UNTERSUCHUNGEN ZUR SUBJEKTIVEN BEWERTUNG DES SCHWEREGRADES VON PSYCHIATRISCHEN UND SOMATISCHEN ERKRANKUNGEN

INAUGURALDISSERTATION

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Psychologie
der Johann Wolfgang Goethe-Universität
zu Franfurt am Main

vorgelegt von

Hans J. Wollny

aus Wasserburg/Inn

1984

UNIV.
LONDON
LIBR.

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41538120_0009

UNTERSUCHUNGEN ZUR SUBJEKTIVEN BEWERTUNG DES SCHWEREGRADES VON PSYCHIATRISCHEN UND SOMATISCHEN ERKRANKUNGEN

INAUGURALDISSERTATION

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Psychologie
der Johann Wolfgang Goethe-Universität
zu Franfurt am Main

vorgelegt von

Hans J. Wollny

aus Wasserburg/Inn

1984

1. Gutachter: Prof. Dr. F. Süllwold

2. Gutachter: Prof. Dr. W. Lauterbach

3. Gutachter: Prof. Dr. L. Schmidt

Tag der mündlichen Prüfung: 4. 7. 1984

Inhaltsverzeichnis

		Seite
1	Einleitung: Der Hintergrund der Untersuchung und die Fragestellung	1
2	Empirische Untersuchungen zur Einordnung von Krankheiten	8
3	Die Methodik der Untersuchung	19
3.1	Die Wahl der Skalierungsmodelle	19
3.1.1	Der pragmatische Gesichtspunkt bei der Auswahl des Skalierungsverfahrens	19
3.1.2	Die meßtheoretische und modelltheoretische Legitimation des Skalierungsverfahrens	21
3.1.3	Das Prinzip der Methode des Paarvergleichs	31
3.1.4	Rechenverfahren und Prüfung der Skalierungshypothesen am Beispiel der Gruppe der Medizinstudenten (Med 1)	37
3.1.4.1	Prüfung der individuellen Urteilskonsistenz	37
3.1.4.2	Die Berechnungsverfahren nach Fall V und Fall III und die Prüfung der Güte der Anpassung	44
3.1.4.3	Das Iterationsverfahren nach GULLIKSEN und die Indices der Skalierbarkeit	51
3.2	Die Auswahl der Krankheitsbilder und die Konstruktion der Paarvergleichsliste	55
3.2.1	Exkurs: Zum Problem der Anwendung des Krankheitsbegriffes auf psychische Phänomene	55
3.2.2	Auswahl der Krankheitsbilder	64

Seite

3.2.3	Erstellung des Untersuchungsinstrumentes zum Paarvergleich und Modus der Vorgabe	67
3.3	Bedeutung des Informationsstandes über die Krankheiten und der persönlichen Kriterien für die Beurteilung	71
3.3.1	Informiertheit über die Krankheitsbilder	71
3.3.1.1	Kursus in Psychiatrie für die Medizinstudenten	71
3.3.1.2	"Wissenstest" über Krankheiten	73
3.3.2	Erfassung der Kriterien der Beurteilung	75
3.4	Kennzeichnung der Beurteilergruppen und Schema der Untersuchung	76
3.4.1	Charakteristik der Untersuchungsgruppen	76
3.4.2	Schema der kontrollierten und variierten Einflußgrößen	84
4	Ergebnisse	87
4.1	Die Skalierung durch die Gruppen der Medizinstudenten (Med 1 und Med 2)	87
4.2	Die Kriterien der Medizinstudenten für die Einstufung der Krankheiten	1o4
4.3	Die Patientengruppen als Beurteiler	11o
4.4	Skalierung der Krankheiten und Informiertheit bei den Patientengruppen	115
4.5	Kriterien der Patientengruppen für die Einstufung der Krankheiten	125
4.6	Ergebnisse der Kontrollgruppen	13o
4.7	Informationsstand über die Krankheiten und Kriterien der Einstufung in Relation zur Skalierung bei den Kontrollgruppen	134

		Seite
4.8	Prüfung auf geschlechtsspezifische Unterschiede	137
4.9	Der Vergleich aller Gruppen gegeneinander: Die differierende Bewertung körperlicher und psychiatrischer Krankheitsbilder	145
4.1o	Psychologische Charakteristik typischer Beurteilergruppen und Beurteiler: Kontrastgruppenvergleich und kontrastierend urteilende Individuen	157
4.1o.1	Kontrastierende Gruppen	157
4.1o.2	Kontrastierend urteilende Einzelne	174
5	Diskussion der Ergebnisse	178
5.1	Kritische Bewertung der Methode	18o
5.1.1	Die Paarvergleichstechnik und die spezifische mathematisch-statistische Auswertung	18o
5.1.2	Kritische Bewertung der Ergebnisse im Licht des Untersuchungsplanes	184
5.2	Die Urteilsunsicherheit über die psychiatrischen Erkrankungen: Überlegungen zu psychologischen und therapeutischen Konsequenzen	191
6	Zusammenfassende Darstellung	199
6.1	Hintergrund und empirisch bearbeitetes Umfeld	199
6.2	Fragestellung	2oo
6.3	Entwicklung der Untersuchungs- und Auswertungsmethodik	2oo
6.4	Ergebnisse	2o7

Seite

Anhang 1: Literaturverzeichnis 213

Anhang 2: Die Paarvergleichslisten 219

Anhang 3: "Wissenstest" mit Schwierigkeits-
indices der Items 237

Anhang 4: Verzeichnis der beteiligten Gruppen
und Kliniken bzw. Institutionen 245

Verzeichnis der Tabellen

		Seite
1	: Matrix der absoluten Häufigkeiten aus einem Paarvergleich	45
2	: Matrix der z-Werte aus einem Paarvergleich	46
3	: Berechnung der Skalenwerte nach Fall V	47
4	: Gruppe Med 1: Vergleich der Skalenwerte nach den drei Berechnungsverfahren Fall V, Fall III und Iteration	5o
5	: Prüfung der Kompetenz der Beurteiler	77
6	: Charakteristik der Untersuchungsgruppen Medizinstudenten 1 (Med 1) und Medizinstudenten 2 (Med 2)	79
7	: Charakteristik der Untersuchungsgruppen Tuberkulose-Patienten (Tb) und Alkoholkranke (Alk)	8o
8	: Charakteristik der Untersuchungsgruppen Kontrollgruppe 1 und 2	82
9	: Schematische Darstellung des Untersuchungsplanes	86
1o	: Skalenwerte der Gruppe Med 1	88
11	: Skalenwerte der Gruppe Med 2	99
12	: Gesichtspunkte der Entscheidung im Paarvergleich der Krankheiten bei den Beurteilergruppen Med 1 und Med 2	1o9
13	: Skalenwerte der Gruppe Tb	111
14	: Skalenwerte der Gruppe Alk	112
15	: "Wissenstest" über die Krankheiten	117
16	: Ergebnis des "Wissenstests"	12o
17	: Zusammenhang zwischen der kombinierten Variable Bildungsstand (BIST), Anzahl der zirkulären Triaden im Paarvergleich (ZT) und Gesamtrohwert im "Wissenstest" (GRW)	123

Seite

18 : Skalenwerte der Gruppe Kontr 1 131

19 : Skalenwerte der Gruppe Kontr 2 132

2o : Einstufung der Krankheiten durch nach Geschlecht unterschiedene Untergruppen 139

21 : Skalenwerte (Iteration) geordnet entsprechend Kontr 1 146

22 : Interkorrelationsmatrix der Skalen (Iteration) der sechs Beurteilergruppen 153

23 : Interkorrelationsmatrizen der auf 1o (organische Krankheiten) bzw. auf 5 (psychiatrische Krankheiten) Skalenwerte reduzierten Skalen 154

23.1 : Produkt-Moment-Korrelationen r

23.2 : Rangkorrelationen R_S

24 : Kennwerte der Kontrastgruppen 162

25 : Interkorrelationen der Extremgruppen-Skalen 172

25.1 : Extremgruppen untereinander

25.2 : Extremgruppen und ursprüngliche Beurteilergruppen

Verzeichnis der Abbildungen

		Seite
1	: Skalen der sechs Beurteilergruppen	89
2	: Graphische Darstellung der Rangplatzverteilungen der Krankheiten	93
3	: Skalen (Iteration) der identischen Beurteilergruppen aus Med 1 und Med 2	1oo
4	: Skala der Gruppe Kontr 2 und Skalen von Untergruppen aus Kontr 2 (jeweils Iteration)	1o2
5	: Relative Häufigkeiten und Rangfolge der von Med 1 und Med 2 genannten Kriterien für die Beurteilung der Krankheiten	1o5
6	: Skalen aller Beurteilergruppen nach linearer Skalentransformation	114
7	: Skalen der Gruppen Med 1 und Alk, getrennt nach Geschlecht	143
8	: Skalenpositionen der somatischen Krankheiten ohne die psychiatrischen Erkrankungen	148
9	: Skalenpositionen der somatischen Krankheiten ohne die psychiatrischen Erkrankungen	149
1o	: Skalenpositionen der psychiatrischen Erkrankungen ohne die somatischen Krankheiten	15o
11	: Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Schizophrenie	16o
12.1	: Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Neurose	167
12.2	: Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Alkoholismus	168
12.3	: Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Hirnarteriosklerose	169

Seite

12.4 : Skalen kontrastierender Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von (endogene) Depression 17o

12.5 : Skalen kontrastierender Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Schizophrenie 171

1 Einleitung: Der Hintergrund der Untersuchung und die Fragestellung

Man kann sich die Frage stellen, ob es einen über ein eher theoretisches Interesse hinausgehenden Sinn hat, subjektive Einschätzungen des Schweregrades von verschiedenen Krankheiten festzustellen. Man mag eher denken, daß "objektive" Kriterien für den Schweregrad von Krankheiten Bedeutung haben, ähnlich wie man bei der Behandlung eines definierten Krankheitsbildes die therapeutischen Maßnahmen stufen muß nach dem Schweregrad. Es könnte daraus folgen, daß man nach einer solchen "objektiven" Skala etwa den materiellen Gesamtaufwand für Erforschung und Behandlung bestimmter Krankheiten politisch gewichtet.

Daß das faktisch nicht geschieht, muß zu der Überlegung führen, daß es wohl keine allgemein verbindliche objektive Rangordnung gibt. Bei genauerer Betrachtung ist nämlich nicht abzusehen, nach welchem Kriterium oder nach welcher Kombination von Kriterien eine solche objektive Skala erstellt werden sollte.

In persönlichen Anschauungen gibt es aber wohl durchaus differierende, vermutlich aber andererseits weitgehend interindividuell übereinstimmende "Skalen" des Schweregrades. Ein Beispiel für einen solchen Meinungsstreit über den Schweregrad von verschiedenen Krankheiten soll helfen, den Hintergrund und die Konsequenzen dieser Frage näher auszuleuchten:

Auf einer wissenschaftlichen Tagung im Herbst 1980 in Bonn zum Thema "Beiträge der Psychologie zur Schizophrenieforschung" kam es unvermittelt zu einer heftigen Diskussion, als ein teilnehmender Neurologe einwarf, "so schlimm" sei nun die Schizophrenie auch nicht. Es gebe in der Neurologie viel schlimmere Krankheitsbilder, das hätten alle Beteiligten doch wohl schon gesehen. Diese Äußerung rief den erbitterten Widerstand der Mehr-

heit der Versammlung hervor. Dem Diskutanten wurde eine Fülle von teilweise sehr emotionalen Argumenten entgegengehalten, von denen zwei im Gedächtnis geblieben sind: Er kenne die Erkrankung, besonders aber das Schicksal vieler Erkrankter, nicht hinreichend, und: Er vergleiche Unvergleichbares.

Diese aus dem Gedächtnis zitierte Diskussion war nicht der Anlaß für die hier vorgelegte Untersuchung zum Problem der subjektiven Bewertung des Schweregrades von psychischen Erkrankungen in Relation zu somatischen Erkrankungen. Die Untersuchung war zu diesem Zeitpunkt bereits konzipiert und einige der Untersuchungsgruppen bereits befragt. Die Diskussion zeigte aber das Interesse an der Fragestellung; noch mehr allerdings beleuchtete sie zwei Kernprobleme: Welche Bedeutung hat der Grad der Kenntnis über eine Krankheit für die Einschätzung, als wie schlimm man sie empfindet? Kann man somatische und psychische Erkrankungen daraufhin vergleichen, wie schlimm sie empfunden werden, oder werden psychische Erkrankungen vielleicht als etwas ganz anderes gesehen? Diese zweite Frage ist gleichbedeutend mit der Frage nach der adäquaten Untersuchungsmethodik: Lassen sich somatische und psychische Erkrankungen nach ihrem subjektiv empfundenen "Schlimmheitsgrad" überhaupt auf einer psychologischen Dimension skalieren?

Wie zu vermuten ist und wie diese geschilderte Diskussion verdeutlicht, gibt es ein allgemeines Interesse daran, wie Krankheiten nach dem Grad ihrer Schwere beurteilt werden. Das hängt sicherlich damit zusammen, daß man entsprechend der Schwere einer Krankheit bereit ist, therapeutischen Aufwand zu treiben, auch gesundheitspolitische Folgerungen zu ziehen, sich publizistisch mit einer solchen Krankheit zu befassen (was wiederum Rückwirkungen auf die allgemeine Beurteilung einer Krankheit hat). Sicherlich leiten aber auch die verschiedenen medizinischen Fachdisziplinen einen Teil

ihres Selbstwertgefühles daraus, mit wie schweren Erkrankungen sie befaßt werden. Das dürfte auch erklären, daß eine kontroverse Diskussion sehr schnell starke Emotionen freisetzt.

Aber es ist festzuhalten, daß es jedenfalls keine verbindliche, objektive Skala der Schwere von verschiedenen Erkrankungen gibt. Vielleicht gibt es aber so etwas wie einen recht allgemeinen, impliziten Konsensus, woran sich gleich eine weitere Frage anschließen kann, ob Ärzte (bzw. angehende Ärzte) und (derzeitige wie potentielle) Patienten eigentlich gleiche Überzeugungen haben.

Aus diesen Überlegungen heraus wäre es gerechtfertigt, diese Fragen dem Gebiet der Medizinischen Psychologie zuzuordnen.

Hier ist die Intention der Frage jedoch anders zu sehen. Sie steht auf dem Hintergrund der Klinischen Psychologie, als Teil eines umfassenderen Problemzusammenhangs, der wichtig erscheint und einer wissenschaftlich gesicherten, aber in seinen Detailfragen auch sorgfältig abgeklärten Lösung zugeführt werden sollte. Es ist der Gesamtkomplex der psychotherapeutischen Arbeit mit psychiatrischen, insbesondere psychotischen und hier wiederum schizophrenen Patienten.

Auf dem Hintergrund der nachgewiesenen Behandelbarkeit von psychotischen Erkrankungen, sowohl grundlegend pharmakotherapeutisch als auch weiterführend psychotherapeutisch, zeigt sich vor allem in der praktischen psychotherapeutischen Arbeit ein Kernproblem. Bei Veröffentlichungen über einzelne Methoden oder Schritte einer solchen Therapie wird implizit immer vorausgesetzt, daß der Patient mitarbeitet. In der alltäglichen Arbeit scheitern aber viele Behandlungs- und vor allem psychotherapeutische Stabilisierungsversuche, auch Bemühungen um Bearbeitung von Reststörungen mit psycholo-

gischen Mitteln, weil der Betroffene nicht von Anfang an eine adäquate Einstellung zu seiner Erkrankung findet.

Dies Problem ist auch bei den medizinischen Teildisziplinen, die mit der Behandlung von längerfristigen Erkrankungen zu tun haben, zunehmend in den Blick gekommen. Hierbei geht es um die Frage, warum ein manchmal erheblicher Prozentsatz von Patienten den Behandlungsanweisungen nicht folgt, wobei im Vordergrund der Betrachtung sicherlich die medikamentöse Therapie steht. Arbeiten hierzu, die unter dem in der Sozialpsychologie in einem weiteren Sinn eingeführten Begriff "compliance" (vgl. HILGARD, ATKINSON & ATKINSON 1971 S. 521 f.) eingeordnet werden, haben im amerikanischen Bereich für das Feld der Psychiatrie bereits eine längere Tradition. Man bemühte sich bereits seit längerem, Einflußfaktoren auf "compliance" abzuklären und therapeutische Ansätze zu ihrer Verbesserung vorzuschlagen (Literatur im Übersichtsartikel von HERMANN 1979).

Bei einer für den deutschsprachigen Bereich vorgelegten neueren Arbeit (LINDEN 1979) wird jedoch speziell dem klinischen Psychologen ins Auge springen, daß das Problem der Einstellung des Patienten zu seiner Krankheit, wie auch seine Überzeugungen von seiner Krankheit nur passager erwähnt wird, aber nicht als wesentlicher Faktor gesehen wird. So scheint etwa ein genannter Ansatz "Therapie der Verbesserung von Compliance" fast ausschließlich als Management äußerer Kontingenzen gesehen zu werden. Ohne die Bedeutung solcher äußeren Kontingenzen bestreiten zu wollen, läßt sich aus psychologischer Sicht begründet vermuten, daß ein ganz wesentlicher Faktor zumal bei psychischen Erkrankungen ist, was eben der Patient über seine Krankheit denkt, ob er zumal bei psychiatrischen Erkrankungen sich überhaupt als krank sieht, weswegen er diese aufgrund rein äußerer Umstände eigentlich als unumgehbar erscheinen-

de Einsicht dennoch möglicherweise massiv abwehrt.

Dieser letzte Problemkreis wird im psychiatrischen Bereich traditionell als "Krankheitseinsicht" angesprochen. Überraschenderweise wird in etlichen klassischen, aber auch in heute verbreiteten psychiatrischen Lehrbüchern auf diese Probleme überhaupt nicht eingegangen (z.B. BLEULER 1969, WIECK 1977, HUBER 1974). Wenn Autoren sich mit diesem Problem auseinandersetzen, lassen sich zwei Grundpositionen unterscheiden, die allerdings nicht kontradiktorisch, sondern gewöhnlich nur konträr gegeneinander gestellt sind:

- Zum Wesen jeder oder aber wenigstens einiger psychiatrischer Krankheitsbilder bzw. seelischer Erkrankungen gehört per definitionem die mangelnde Krankheitseinsicht (z.B. PAULEIKHOFF & MESTER 1973).
- Mangelnde Krankheitseinsicht ist ein Problem, das mehr oder weniger häufig zur psychiatrischen bzw. seelischen Erkrankung kontingent ist (z.B. JASPERS 1959, WEITBRECHT 1968).

Neben den vielen empirischen Arbeiten zur Pharmakotherapie besonders der Psychosen liegt bereits eine stattliche Zahl von Untersuchungen, Modellen und Vorschlägen zur ergänzenden psychologischen Therapie und zur Soziotherapie dieser Erkrankungen vor. Für das Stadium zwischen Aufsuchen des Arztes bzw. (oft unfreiwilliger) Einweisung in die Klinik und Einsetzen eigener Mitarbeit bei der Behandlung (soweit überhaupt erreicht) besteht aber eine empirische Lücke. Dabei ist festzustellen, daß gerade dieses Anfangsstadium im klinischen Alltag erhebliche Schwierigkeiten bereitet. Deutlich wird das ganz besonders dadurch, daß die - in anderen Bereichen selbstverständliche - Frage eines Patienten nach seiner Diagnose große Unsicherheit auslösen kann, wie man antworten soll.

So ist die Frage, wie man durch äußere Maßnahmen die Mitarbeit des Patienten bei seiner Behandlung verbes-

sern, steuern, besser kontrollieren kann, weniger grundlegend. Wesentlicher ist die Frage danach, wie man eine adäquate Einsicht eines Betroffenen in seine Situation als Kranker fördern kann, wie evtl. "Abwehr" überwunden werden kann, wie man helfen kann, das Erleben der Krankheit zu verarbeiten, wie man zur eigenen Bemühung motivieren kann. Hierzu ist man auf psychologisch-psychotherapeutische Mittel verwiesen.

Dies ist nun der Gesamtzusammenhang, den man bezeichnen könnte als "das Problem der Krankheitseinsicht und der Einstellung zur Krankheit, Möglichkeiten einer therapeutischen Beeinflussung und Effektivität solcher Maßnahmen für die eigentliche Pharmako-, Psycho- und Soziotherapie".

Als erste grundlegende Frage interessiert in diesem Zusammenhang, wie eigentlich psychiatrische Erkrankungen in Relation zu körperlichen Erkrankungen von potentiell Betroffenen gesehen werden. Es wurde bereits begründet, daß es eine "objektive" Skala für die Schwere von Krankheiten nicht gibt und auch nicht geben kann. Andererseits läßt sich vermuten, daß die Bereitschaft zur Verleugnung von Krankheiten ("Abwehr") mit der subjektiv empfundenen Schwere oder Bedrohlichkeit oder "Schlimmheit" zusammenhängen könnte, wobei eine subjektive Skala der "Schlimmheit" zu postulieren ist.

Die erste konkrete Frage dieser Untersuchung ist, ob sich eine solche Skala auf der Grundlage eines Gruppenurteils erstellen läßt. Dabei ist zugleich zu prüfen, ob unter einem gemeinsamen, vorherrschenden Aspekt geurteilt wird. Denkbar ist nämlich auch, daß von vornherein differierende Gesichtspunkte einfließen, etwa derart, daß einige Personen zumal bei psychischen Erkrankungen soziale Aspekte absolut in den Vordergrund stellen, andere ausschließlich den Aspekt der körperlichen Beeinträchtigung oder der Lebensgefahr o.ä. als entscheidend für die Beurteilung subjektiv empfundener

Schwere ansehen. Das bedeutete dann, daß es nicht eine einzige psychologische Dimension "Schlimmheit" gäbe, sondern etwa zwei Dimensionen, "sozial bedrohlich" und "lebensbedrohlich", oder noch mehr als zwei.

Neben dieser zweiten Frage nach der Dimensionalität der zu suchenden Skala oder Skalen ergibt sich als Drittes eine Gruppe von Fragen: Welche Faktoren beeinflussen die Einstufung von Krankheiten in welcher Weise? Denkbar sind Unterschiede, welche korreliert sind mit Unterschieden nach Alter, Geschlecht, Bildungsstand, Grad detaillierterer Kenntnis, eigener Erfahrung. Hier gilt es also, typische Gruppen zu finden, die sich in definierten Kenngrößen voneinander unterscheiden, um sie gegeneinander vergleichen zu können.

2 Empirische Untersuchungen zur Einordnung von Krankheiten

Es gibt nur eine geringe Anzahl empirischer Untersuchungen, die Krankheiten miteinander vergleichen; zumeist unter dem Kriterium, vor welcher Krankheit man mehr Angst habe, welche Krankheit man für schwerer halte. Dabei zeigen sich folgende sachliche Zusammenhänge, in welche diese Untersuchungen einzuordnen sind: Einmal empirische Untersuchungen zu Maßnahmen der Gesundheitsvorsorge (also ein gesundheitspolitischer Kontext), in den USA zeitlich zuerst in Zusammenhang mit den Impfkampagnen gegen Kinderlähmung, dann bei allen Maßnahmen im Dienst zur Früherkennung von Krebs. Zum Zweiten zwei Untersuchungen (eigentlich eine Untersuchung mit einer Fortsetzung) einer Autorengruppe zu Fragen der Rehabilitation von Körperbehinderten. Schließlich findet sich im deutschsprachigen Bereich eine Arbeit, die nicht Krankheiten, aber Beschwerden skalieren läßt.

Die erste von diesen empirischen Arbeiten, die wir auffinden konnten, stammt von ROBBINS (1962). Er will versuchen, die psychologischen Faktoren näher zu erhellen, welche Menschen davon abhalten, die Dienste des "Health Service" zur Früherkennung von Krankheiten in Anspruch zu nehmen. Er bezeichnet seine Studie als "exploratorisch". So besteht der erste, wohl wesentlichere Teil aus einem hochstrukturierten Interview, das einmal klären will, auf welches von neun vorgegebenen Krankheitsbildern (Krebs, Herzkrankheit, Kinderlähmung, Geisteskrankheit, Tuberkulose, Magenbeschwerden, Rückenbeschwerden, Zahnbeschwerden, Erkältung) sich vorzugsweise die Befürchtungen und Ängste richten, zum anderen, welcher inhaltliche Aspekt der Angst im Vordergrund steht. Dabei sind neun inhaltliche Aspekte (diffuse Angst, Schmerz, Dauer/Heilbarkeit, Behandlungsform, Veränderung der eigenen Person, Beeinträchtigung der Leistungsfähigkeit, soziale Folgen, finan-

zielle Folgen, Erfahrung durch Erkrankung von Angehörigen oder Freunden) im voraus festgelegt. Sie geben das Gerüst für das Interview und werden in Bezug auf jede Krankheit abgefragt, nachdem vorher erfragt wurde, welche Krankheit aktuell im Vordergrund der Besorgnisse steht ("Saliency"). Weiterhin wurde eine Einstufung der Krankheiten erbeten "im Vergleich zu den meisten Krankheiten". Interviewt wurden 46 Mütter, die eine Gesundheits-Beratungsstelle für Kinder aufgesucht hatten; sie waren diejenigen von insgesamt 96 Müttern, die einen zugesandten Angstfragebogen (den anderen Teil der Untersuchung) über Krankheiten beantwortet hatten.

Eine Auswertung der Interviewdaten (Rating; Prozentzahlen der Zuordnung zu inhaltlichen Gesichtspunkten, etwa Bezeichnung als "sehr ernst") ergab eine Rangreihe der Schwere, angeführt von Krebs, gefolgt von Herzkrankheit, dann Geisteskrankheit, dann Kinderlähmung und Tuberkulose, dann die restlichen. Während Krebs wohl als ernsteste Krankheit wie auch als diejenige, mit der man sich (in Gedanken) am meisten befaßt und zugleich darüber sorgt, angesehen wurde, findet sich keine Übereinstimmung der Rangordnung in Bezug auf diese drei Kategorien (Saliency, Seriousness, Source of experienced concern) bei den anderen Erkrankungen.

Die Ergebnisse des Interviews werden im zweiten Teil der Untersuchung in Beziehung gesetzt zu Ergebnissen der Fragebogenbeantwortung. Dieser Fragebogen, dessen Reliabilität eher provisorisch vorher an einer Studentengruppe (N=52) und einer Gruppe von Angestellten eines Sozialdienstes (Social security administration; N=39) geprüft wurde, wurde entsprechend den Gesichtspunkten des Interviews aufgeteilt, einmal in Subskalen, nach Krankheitsbild, zum anderen in Subskalen nach den oben genannten inhaltlichen Kategorien. Auf der Subskalenebene zeigten sich keine Zusammenhänge zu den Interviewdaten. Eine Faktorenanalyse des Fragebogens

(oblique Rotation; es wird nicht angegeben, inwieweit die Gesamtvarianz durch die extrahierten Faktoren aufgeklärt wird) erbrachte neben schwer interpretierbaren oder nur auf eine Krankheit bezogenen (spezifischen) Faktoren zwei, bei denen sich nach Bildung von Faktorenscores aus den Fragebogen Zusammenhänge zwischen im Interview geäußerter Angst vor ernsten Krankheiten und dem Faktorenscore prüfen ließen.
Für einen Fragebogenfaktor, der als affektiv-emotional getönte Angst beschrieben wird, zeigte sich ein signifikanter Zusammenhang zur geäußerten Angst im Interview, andererseits kein Zusammenhang zwischen dieser und einem Fragebogenfaktor, der als eher kognitiv, als "Denken an mögliche Beeinträchtigung durch Krankheiten" beschrieben wird. Der "emotionale" Angst-Faktorscore (Fragebogen) korreliert gleich signifikant mit der Angst-Skala von TAYLOR (M.A.S.). Es zeigen sich Zusammenhänge zwischen emotional getragener Angst und niedrigerem Bildungs- und Berufsniveau, weiblichem Geschlecht, höherem Alter und höherer Kinderzahl. Während der emotionale Angst-Faktor keinen Zusammenhang zeigte mit Erfahrungen mit ernsten Krankheiten im engeren persönlichen Umfeld, ergab sich ein gegensinniger Zusammenhang zum sogenannten "kognitiven" Faktor. Was sich nicht zeigte, war ein Zusammenhang zwischen Maßen des Fragebogens und Gesundheitsverhalten, wie z.B. Teilnahme an einer Grippe-Schutzimpfung.

Ebenfalls im Zusammenhang mit der vorbeugenden Gesundheitsfürsorge entstand die Arbeit von LEVINE (1962), in der Umfragedaten von insgesamt 297o Probanden aus 3o Verwaltungsbezirken der USA über sechs Krankheiten, Krebs, Tuberkulose, Schlaganfall, Arthritis, Geburtsschäden und Kinderlähmung (Poliomyelitis) erhoben wurden.

Die Probanden hatten jede Krankheit auf einer 5-Punkte-Skala nach dem Maß ihrer Angst vor der Krankheit einzu-

stufen. Auf einer umgerechneten Skala (o - 1oo) ergaben sich folgende mittleren Angst-Skalen-Werte: Krebs = 72, Polio = 58, Schlaganfall = 52, Arthritis = 5o, Geburtsdefekte = 46, Tuberkulose = 38. Als Korrelate der Angst vor einer bestimmten Krankheit wurden aus den Ergebnissen identifiziert das Niveau der Schulbildung (gegensinniger Zusammenhang: niedrige Schulbildung - hohe Angstwerte), das berufliche Niveau, nähere Kenntnis eines Krankheitsfalles, und damit im Zusammenhang selbsteingeschätztes genaueres Wissen über eine Krankheit, dann Lebensalter in dem Sinne, daß die Angst vor für das eigene Lebensalter wahrscheinlichen Krankheiten hoch ist. Hierbei stellt Krebs allerdings eine Ausnahme dar. Die Angst vor Krebs ist in ihrem Ausmaß vom Alter unabhängig. Die engste Beziehung besteht zwischen Ausmaß der Angst vor einer Krankheit und Überzeugung vom Maß ihrer Häufigkeit bzw. Verbreitung. Weiterhin zeigte sich eine Beziehung zwischen Maß der Angst und geringerer Erreichbarkeit medizinischer Institutionen.

Aus diesen Ergebnissen wird der Schluß gezogen, daß Angst vor einer Krankheit vorzugsweise auf rationalen, die eigene Person in Betracht ziehenden Beweggründen basiert.

Auch die dritte zu berichtende Arbeit steht im Zusammenhang mit der öffentlichen Gesundheitsfürsorge. Es ist eine Untersuchung von JENKINS & ZYZANSKI (1968), die mit einem eigens entwickelten semantischen Differential die Überzeugungen über drei Krankheiten, Polio, Krebs und Geisteskrankheit, an einer Stichprobe von 436 Erwachsenen untersucht.

Die insgesamt 16 Skalen des semantischen Differentials, die u.a. Vorstellungen über Häufigkeit, Schmerzhaftigkeit, Schwere, Häufigkeit des Denkens an die Krankheit, des Sprechens darüber abfragten, wurden für jede Krankheit getrennt korreliert und faktorenanalysiert. Die

Autoren versuchen, Gemeinsamkeiten zwischen den Faktorenstrukturen herauszustellen. Sie sehen diese Gemeinsamkeiten darin, daß bei allen Krankheitsbildern, allerdings mit den geringsten Anteilen an der aufgeklärten Varianz, jeweils zwei Faktoren zu finden sind. Den einen interpretieren die Autoren als "human mastery" (etwa: Beherrschbarkeit der Krankheit durch Kenntnis der Ursachen und daraus folgender Therapierbarkeit), den anderen als "soziale Bewertung" (social acceptability). Polio ist dann noch gekennzeichnet durch einen Faktor "persönliche Affiziertheit" (die Wahrscheinlichkeit, selber zu erkranken, wird hoch eingeschätzt, und man denkt oft daran und spricht davon) und "Akutheit, Schwere" (schwer, schnell fortschreitend, schmerzhaft). Krebs ist charakterisiert durch einen Faktor "Furcht" und einen weiteren "Salience" (man denkt oft daran, spricht oft davon). Geisteskrankheit hat in dieser Untersuchung nur eine Struktur aus drei Faktoren, deren erster wiederum als "persönliche Affiziertheit" wie bei Polio gesehen wird.

Die Autoren versuchen, unter dem Aspekt der Beteiligung möglichst vieler potentieller Patienten an einem Gesundheitsvorsorgeprogramm Schlußfolgerungen zu ziehen: Die besten Aussichten für ein Ansprechen eines solchen Programmes sehen sie, wenn die Zielgruppe eine hohe persönliche Affiziertheit zeigt, wenn der Krankheit die soziale Stigmatisierung (die nach der Untersuchung bei Polio und Geisteskrankheit noch besteht) genommen wird und vor allem, wenn im öffentlichen Bewußtsein die Meinung verbreitet ist, daß die Ursachen der Erkrankung geklärt sind und sie damit beherrschbar ist.

Im eingegrenzteren Zusammenhang mit den Bemühungen um die Vorsorgeuntersuchung bei Krebs entstand die Untersuchung von KNOPF (1976) in Lancaster (Großbritannien). Einer repräsentativen Stichprobe von 75o Frauen wurden

neben speziellen Fragen zur Krebsvorsorge auch Fragen vorgelegt, bei denen sie mehrere Krankheiten nach verschiedenen Gesichtspunkten einordnen sollten: Bedeutung als Todesursache, größte persönliche Besorgtheit, Heilbarkeit.

Zwei Drittel der Befragten sahen Krebs (wie in einer früheren Untersuchung 1966; Daten der referierten Untersuchung von 1973) als häufigste Todesursache an. Nur die Gruppe aus der höchsten von drei sozialen Schichten wich dahingehend ab, daß sie mehr als die anderen beiden Gruppen den tatsächlichen Häufigkeiten (wobei zu fragen ist, wie reliabel die dokumentierten Todesursachen wirklich sind) näherkam und kardio-vaskuläre Erkrankungen als häufigste Todesursache nannte (28 %; mehr als 1966). Krebs führte auch die Rangreihe der Krankheiten an, die die größte persönliche Besorgtheit erregten, 1973 ein wenig geringer als 1966 (52.5 zu 55.5 % Nennungen). Die Rangreihe der prozentualen Nennungen war für beide Untersuchungszeitpunkte gleich: Krebs, psychische Krankheit, Herzkrankheit, Bronchitis, Tuberkulose. Wiederum in der oberen Sozialschicht und bei jüngeren Frauen zeigte sich jedoch ein Anstieg der Nennungen für psychische Krankheit (mental disease) als am meisten persönliche Besorgtheit auslösend von 22.6 % auf 27.9 %. Während weniger Befragte Krebs für unheilbar halten als vor 6 Jahren, denken ein wenig mehr Frauen bei der zweiten Befragung, daß Geisteskrankheit unheilbar sei (17.1 zu 21.2 %). Krebs wird von weniger Frauen für gewöhnlich als heilbar angesehen als psychische Krankheit, von deutlich mehr jedoch gegenüber psychischer Krankheit als "manchmal heilbar". Von den meisten wird Tuberkulose als "gewöhnlich heilbar" angesehen.

Einer frühzeitigen Behandlung räumt zu beiden Zeitpunkten ein gleich großer Prozentsatz der Befragten für Krebs einen positiven Wert ein. Für Geisteskrankheit

ist dieser Prozentsatz, anders als bei allen abgefragten Krankheitsbildern, deutlich abgesunken (von 83 auf 71.3 %).

Die Autorin zieht die sehr wesentlich erscheinende Schlußfolgerung, daß eine öffentliche Publizität einer Krankheit nicht unbedingt Verminderung von Pessimismus oder Furcht bewirkt und so früher Intervention möglich macht. Im Zeitraum zwischen beiden Untersuchungen war in Großbritannien die Häufigkeit publizistischer Behandlung beider Krankheiten gewachsen. Während aber die Beschäftigung mit Krebs (nach dem Referat der Autorin) eher ein kohärentes, geschlossenes Verständnis nahelegte, standen vor allem bei Presseveröffentlichungen über psychische Krankheit die Häufigkeit neurotischer Störungen in der Gesamtbevölkerung, Skandale in Nervenkliniken und Kontroversen über die richtige Behandlungsmethode im Vordergrund.

Alle diese drei Untersuchungen, die in einem gemeinsamen Problemzusammenhang stehen, aber methodisch sehr unterschiedlich angesetzt sind, zeigen ein gleiches Ergebnis. Unter allen Aspekten wird Krebs als schwerste, am meisten ängstigende, als schlimmste Krankheit bewertet. Geisteskrankheit oder psychische Krankheit scheint zu den schwereren gerechnet zu werden. Ein Problem liegt in der Begrifflichkeit der Bezeichnungen "mental illness", "mental disease". Im deutschsprachigen Bereich ist der Begriff Geisteskrankheit relativ eindeutig; damit ist umgangssprachlich gewöhnlich Schizophrenie gemeint, während psychische Krankheit ein umfassender Begriff ist. Weiterhin sind dann noch die diagnostischen Differenzen zwischen dem angloamerikanischen Gebiet und dem Gebiet deutschsprachiger Psychiatrie zu bedenken.

Zu den bisher referierten Arbeiten aus dem angloamerikanischen Bereich ist eine deutschsprachige Publikation zu ergänzen, die ebenfalls im Zusammenhang mit öffent-

licher Gesundheitsfürsorge entstanden ist. Die Publikation von SCHÖNROK (1978) hat aber vor allem methodisches Interesse. Es soll ein Verfahren zur Erfassung des "Image" einer Krankheit vorgestellt werden, das längerdauernde Einstellungen erfaßt, von denen wiederum eine engere Beziehung zum tatsächlichen Verhalten angenommen wird.

Den Probanden wurden pro Krankheit 6 Kriterien vorgegeben, die jeweils 4-stufig skaliert werden konnten. So wird dann jeweils ein Kriterienprofil der Gesamt-Untersuchungsgruppe oder von Einzelgruppen von Probanden für jede Krankheit erstellt, die einen Vergleich verschieden geordneter Probandengruppen zulassen. Die Bedeutung des Verfahrens, vor allem die psychometrische Qualität ist eher kritisch zu bewerten. Für uns ist immerhin von Interesse, daß sich drei Gruppen in Bezug auf die Gefährlichkeit ergeben: Krebs und Herzinfarkt als sehr gefährlich; Nierenentzündung, Magengeschwür, Gallensteine, Asthma, Kreislaufstörungen, Bluthochdruck und Zuckerkrankheit im Mittelbereich, grippaler Infekt und Krampfadern als wenig gefährlich. Psychische Erkrankungen sind nicht berücksichtigt. Wiederholungsstabil in hohem Maße ist die Beurteilung von Krebs in Bezug auf Gefährlichkeit (r_{tt} = .95), aber auch die Beurteilung dieser Krankheit in Bezug auf andere Kriterien, während die Dauerhaftigkeit der Beurteilung der anderen Krankheiten dagegen deutlich abfällt. Zu bedenken ist, daß Kennwerte für Zeitstabilität aber sowohl von der Dauerhaftigkeit eines erfaßten Merkmals als auch von der Qualität des Untersuchungsinstruments abhängen.

Der zweite sachliche Zusammenhang, in dem nach Einstufungen von Krankheiten oder Behinderungen gefragt wurde, ist die Rehabilitationsforschung.

Vertreten ist er aber nur durch zwei kleinere Berichte einer Arbeitsgruppe WILSON, BEATTY & FRUMKIN (1967)

und WILSON, SAKATA & FRUMKIN (1968), die in ihrer ersten Studie insgesamt 32 Studenten (15 männliche, 17 weibliche) einer Rehabilitationsausbildung nach ihrer persönlichen Einschätzung der Beeinträchtigung durch folgende 8 Behinderungen gefragt haben: Arm-Amputation, Bein-Amputation, Blindheit, unheilbare Herzkrankheit, vollständige Taubheit, schwere Migräne, unheilbare Lungentuberkulose, unheilbare Magengeschwüre.

Verglichen wurden Mittelwerte aus den individuellen Rangplätzen zuerst in der Studentengruppe, getrennt für beide Geschlechter. Blindheit wurde dabei übereinstimmend als am meisten beeinträchtigend eingestuft, dann Migräne. Wesentlich unterschieden sich die Geschlechter bei der Einstufung einer Arm-Amputation: die Frauen betrachteten sie als deutlich schlimmer. Ergänzend befragten die Untersucher in ihrer zweiten Studie eine Gruppe von hochbegabten Erwachsenen (N=5o) sowie einige Professoren des Ausbildungsinstitutes für die Rehabilitationsausbildung. Die Übereinstimmung innerhalb der Gruppen (Konkordanzkoeffizient W) war signifikant und mittelhoch; die Begabten-Gruppe und die Studenten stimmten in ihrer Beurteilung deutlich überein, von ihnen differierte die Professorengruppe. Schwierig erscheint die von den Autoren versuchte Einordnung in einen Zusammenhang mit der Ausbildung als Reha-Berater. Von Interesse ist eher der Versuch einer Einstufung an sich.

Eine dritte Variation der Fragestellung liegt einer Untersuchung von BINDER & ANGST (1981) zugrunde. Diese Untersuchung fragt nicht nach der Skalierung von Krankheiten, sondern läßt von Probanden mit psychosomatischen und neurotischen Störungen Beschwerden skalieren. Die in dieser Untersuchung befragten Probanden waren nach einer Erstuntersuchung mit einem Symptomfragebogen so ausgewählt, daß zwei Drittel über einem cut-off-Wert lagen, der psychosomatische Störungen anzeigt, ein Drittel unter diesem cut-off. Es handelt sich um jeweils

3oo Männer und Frauen im Alter von 2o bis 21 Jahren. Sie wurden in einem strukturierten Interview nach dem aktuellen Vorhandensein von 25 psychosomatischen bzw. neurotischen Syndromen befragt. Bei Vorhandensein wurden sie gebeten, die "subjektive Belastung" auf einer Ratingskala einzustufen. Weiterhin wurde die Häufigkeit sozialer Konsequenzen auf einer Prozentskala dargestellt, wobei die abgefragten Kriterien für soziale Belastung von den Untersuchern vorgegeben wurden: Behinderung bei der Arbeit, Beeinträchtigung der Freizeit, Beeinträchtigung sozialer Kontakte. Schließlich wurden zwei zusätzliche Skalen, "Häufigkeit von Chronifizierungen" und "Häufigkeit einer ärztlichen Behandlung" für die erfragten Syndrome erstellt.

Im Hinblick auf subjektive Belastung wird von Männern wie Frauen zuerst "Depression" (von den Untersuchern als Syndrom definiert) und dann "Angst" als am meisten belastend eingestuft (neben heterogenen Sammelkategorien wie "soziale" und "psychologische Probleme"). Nach BINDER & ANGST (1981) lassen sich weiterhin die Syndrome in zwei Häufungen unterteilen, und zwar die belastenderen mit psychischen Symptomen und die weniger belastenden mit körperlichen Symptomen wie Kopfschmerzen, Schlafstörungen, Rückenbeschwerden, Herzbeschwerden, Magenbeschwerden etc. (aber auch Zwangssymptome). Aus unserer Sicht soll noch herausgehoben werden, daß Männer Alkoholabusus in Bezug auf "subjektive Belastung" bei Null einstufen, bei den Frauen kommt das Syndrom in der Skala nicht vor (ebenso wie Drogenabusus), da nur bei einem Prozent eruiert.

Im Hinblick auf die Beeinträchtigung im sozialen Bereich (Häufigkeitsskalen) liegt Depression bei Männern und Frauen immer im oberen Skalenbereich, Alkoholabusus im unteren Skalenbereich. Hingegen wird Depression weniger häufig als chronifiziert beurteilt (Männer und Frauen etwa zu 7o %) als etwa Kopfschmerzen, Allergien,

Phobien, und bei Frauen phobische Syndrome (mit etwa 9o % hier führend), bei Männern Zwangssyndrome (mit etwa 85 % führend). In Bezug auf die Häufigkeit der Behandlung liegt wiederum Alkoholabusus (bei den Männern) bei o Prozent, Depression jetzt am unteren Ende der Häufigkeitsskala bei Männern und Frauen (etwa bei 15 %). Führend in der Häufigkeit der Behandlung sind bei Männern und Frauen die Sammelkategorie somatische Beschwerden, dann folgen die spezifizierten körperlichen Syndrome.

BINDER & ANGST bezeichnen als wesentliches Ergebnis, daß die (psychischen) Syndrome, welche als subjektiv sehr belastend eingestuft und häufiger als beeinträchtigend genannt wurden, nicht diejenigen sind, welche am häufigsten ärztlich behandelt werden. Das sind die organbezogenen Syndrome.

3 Die Methodik der Untersuchung

3.1 Die Wahl der Skalierungsmethode

Die erste Frage ist: Wie ordnen verschiedene Gruppen von Beurteilern verschiedene Krankheiten nach einem vorzugebenden Kriterium ("psychologische Dimension") an? Diese Frage führt methodisch auf ein Skalierungsproblem.

3.1.1 Der pragmatische Gesichtspunkt bei der Auswahl des Skalierungsverfahrens

Skalierungsprobleme sind im Gebiet der Psychologie im wesentlichen in drei Gebieten bearbeitet worden, klassisch in der Psychophysik seit FECHNER, dann vor allem seit THURSTONE in der Sozialpsychologie, insbesondere im Bereich der Einstellungsforschung. Aber auch die psychologische Diagnostik befaßt sich mit Skalierungsfragen.

Es werden verschiedene Einteilungsgesichtspunkte für psychologische Skalierungsmethoden (abgesehen von dem anderen Aspekt des Skalenniveaus) vorgeschlagen. Dabei ist zu beachten, daß mit gleichen Begriffen nicht immer präzis der gleiche Inhalt gemeint ist (EDWARDS 1957, SIXTL 1967 u. 1982, ORTH 1974, GIGERENZER 1981). SIXTL (1967) hat in der Einleitung seiner Monographie über psychologische Meßmethoden wesentliche Einteilungsgesichtspunkte aufgeführt. Sein Schema ist nicht in allem schlüssig, doch läßt sich danach der pragmatische Aspekt für die Auswahl eines konkreten Verfahrens gut darstellen.

Da hier nicht zuerst die persönliche Einstellung eines Individuums interessiert, entfallen alle Methoden bzw. die ergänzenden Schritte mancher Methoden, welche zu den "Response"- (den "subjektzentrierten") Methoden gezählt werden. Diese kann man summarisch dahingehend charakterisieren, daß mit ihnen der individuelle Punkt

eines Probanden auf einer psychologischen Skala bestimmt werden soll, welche den Grad einer positiven oder negativen Einstellung zu einem Objekt bezeichnet, oder auch die Ausprägung einer Eigenschaft o.ä. Das ist gewöhnlich die Frage der Einstellungsmessung, aber auch der psychologischen Diagnostik. Damit erübrigen sich zugleich die Methoden, die sozusagen in einem Untersuchungsgang (quasi von rückwärts) die Konstruktion einer Skala gestatten u n d die Messung einer individuellen Position. Sie werden auch "reaktionszentrierte" Verfahren genannt.

Die Fragestellung führt also zur Stimulus-Skalierung, zu reizzentrierten Verfahren: Die Krankheiten als Reize sollen skaliert werden; dabei getrennt für jede Gruppe.

Es bleiben zwei Kategorien von Unterscheidungsmerkmalen offen: einmal die Wahl zwischen direkten und indirekten Methoden, dann die Entscheidung zwischen (linearer) eindimensionaler und mehrdimensionaler Skalierung.

Die Entscheidung zwischen eindimensionaler und mehrdimensionaler Skalierung ist im Prinzip a priori zu treffen: Das Kriterium, nach welchem die Beurteiler die Krankheiten einordnen sollen, soll so gewählt werden, daß sich die Krankheiten auf einer einzigen Dimension skalieren lassen. Ob das dann durch die Beurteiler tatsächlich so vollzogen wird, muß empirisch geprüft werden.

Der wesentliche Unterschied zwischen einer direkten und einer indirekten Methode besteht darin, daß bei der direkten Methode der einzelne Beurteiler die Vorstellung einer zumindest ordinalen Skala psychologisch präsent haben muß. Bei der indirekten Methode werden hingegen einfachere Reaktionen verlangt. Die Skalierung erfolgt erst durch Anwendung und gleichzeitige Prüfung definierter Verteilungsannahmen über diese Reaktionen. Eine Entscheidung zwischen beiden Methoden

sollte sich nach den Kriterien einmal der Praktikabilität und zum anderen der Prüfbarkeit von Voraussetzungen richten. Es erscheint eine indirekte Methode einmal praktikabler im Hinblick auf die von den Beurteilern zu fordernde Leistung, zum anderen macht sie weniger empirisch nicht prüfbare Voraussetzungen über das adäquate Urteilsverhalten der Beurteiler.

Als gut praktikables Verfahren bot sich die Paarvergleichsmethode nach THURSTONE (1927) an. Diese Methode ist exemplarisch angewendet worden von SÜLLWOLD (1967) in einer Untersuchung über Gruppenstereotype gegenüber anderen Völkern.

3.1.2 Die meßtheoretische und modelltheoretische Legitimation des Skalierungsverfahrens

Die eher pragmatischen Überlegungen zur Auswahl eines Skalierungsverfahrens sind als erster Schritt dargestellt. Sie stehen auf dem Hintergrund des theoretischen Standards, wie er etwa bei TORGERSON als sogenanntem "Klassiker" entwickelt ist und von SIXTL (1967; 1982) dargestellt wird. Diese Überlegungen stehen deswegen an erster Stelle, weil die tatsächliche Planung der Untersuchung in dieser Weise fortgeschritten ist.

Es ist aber notwendig, sich mit grundsätzlicher Kritik aus neueren Überlegungen zur Theorie der Messung und der Modellbildung in der Psychologie auseinanderzusetzen. Als Verfahren der Wahl, begründet nach pragmatisch-klassischen Kriterien, wurde nach Einteilungsgesichtspunkten eine Skalierungstechnik definiert, die reizzentriert und indirekt ist.

Es ist ein wesentliches Anliegen neuerer Überlegungen, wie sie bei ORTH (1974) und bei GIGERENZER (1981) dargestellt sind, daß eine Datenstruktur bestimmte Voraussetzungen erfüllen muß, ehe sie überhaupt in irgendeiner Weise skalierbar ist. Es ist nicht hinreichend, auf Datenmengen einfach statistische Verfahren anzu-

wenden. Die Sinnhaftigkeit solcher Verfahren und damit ihre Legitimität ist vielmehr zu begründen. Es ist nachzuweisen, daß die Operationen der Beurteiler die Kriterien erfüllen, welche zur Erreichung eines bestimmten, weitere statistische Verarbeitungen rechtfertigenden Skalenniveaus notwendig sind.

An einem Beispiel soll verdeutlicht werden, was damit gemeint ist. In der referierten Untersuchung von LEVINE (1962) erhielten die Beurteiler zur Skalierung von sechs Krankheiten zwei Vorgaben: Zum einen die Definition der Beurteilungsdimension, sinngemäß: "Wie sehr haben Sie Angst vor dieser Krankheit?", zum zweiten die Definition einer Skala mit fünf Punkten (wohl 1 - 5 oder auch o - 4). Es erhält also jede Krankheit (Krebs, Tuberkulose, Schlaganfall, Arthritis, Geburtsschäden und Kinderlähmung) von jedem Beurteiler einen Punktwert für das Maß eigener Angst vor dieser Krankheit.

Die psychologische Analyse zeigt, daß die Beurteiler eine Ordnungsoperation vollzogen haben. Die Erstellung der Größer-Kleiner-Relation, also einer Ordinalskala, geschah durch die Zuordnung von Zahlenwerten zu Objekten (Krankheiten). Dabei implizieren die Zahlenwerte die Ordnungsrelation nach "Größe der Angst".

Nach den Forderungen der Meßtheorie setzt eine Ordnungsstruktur, wie sie im numerischen Relativ (Zahlenwerte als Repräsentation einer Stufung) dargestellt wird, die Erfülltheit von zwei sogenannten Axiomen im empirischen Relativ ("Größe der Angst") voraus: Konnexität (ORTH 1974), oder anders bezeichnet "Vollständigkeit" (GIGERENZER 1981), und Transitivität. Konnexität oder Vollständigkeit meint, daß alle zu beurteilenden Objekte tatsächlich in eine Größenordnung zueinander gebracht sind. Die Transitivität ist dann verletzt, wenn ein Beurteiler zum Beispiel Objekt A als größer ansieht als B, weiterhin B als größer C, dann aber C wiederum als größer A bezeichnet. Diese Anordnung ist widersprüch-

lich; es resultiert ein intransitives Urteil, auch "zirkuläre Triade" genannt. Gefordert wird, daß bei Ordnungsstrukturen zumindest das Axiom der Transitivität empirisch prüfbar sein sollte.

Das ist bei dem von LEVINE (1962) verwendeten, weit verbreiteten Rating-Verfahren nicht möglich. Vielmehr wird durch die Vorgabe der Skala von vornherein die Möglichkeit des Auftretens intransitiver Urteile vermieden. Anders herum formuliert: Es wird gewissermaßen Transitivität beim einzelnen Beurteiler erzwungen. Man kann mit guten Gründen annehmen, daß die Beurteiler in der Untersuchung von LEVINE tatsächlich eine Ordinalskalierung vornehmen. Es ist aber festzuhalten, daß das Ergebnis - zumindest in einem bestimmten Ausmaß - auch artifiziell sein kann. Es ist nicht selten, daß Beurteiler eine Vorgabe, wie sie etwa LEVINE mit der Definition der Beurteilungsdimension "Größe der Angst" vorgibt, spontan differenzieren, etwa in Angst vor dem Sterben, Angst vor bleibenden Behinderungen, Angst vor sozialen Folgen. Wäre das möglich, könnte intransitives Urteilsverhalten resultieren. Daraus folgt die weitere Frage, ob eine Beurteilung etwa solcher Objekte tatsächlich im psychologischen Sinne eindimensional erfolgt, oder ob "Eindimensionalität" sich nicht allein aus der "erzwungenen" Transitivität ergibt.

Kritisch zu beurteilen ist aber die dann von LEVINE (1962) angewandte statistische Operation. Das bezieht sich nicht auf die Transformation der Zahlenwerte von 1 bis 5 in o bis 1oo, weil eine monotone Transformation die ursprüngliche Größer-Kleiner-Ordnung unberührt läßt, also für eine Ordinalskala eine zulässige Transformation ist. Vielmehr setzt der Autor, wie aus der anschließend vorgenommenen Bildung arithmetischer Mittelwerte folgt, über das Ordinalskalenniveau hinausgehend voraus, daß die Beurteiler die vorgegebene Punkte-Skala als Intervallskala mit definierten Abständen angesehen

haben. Die aus diesen statistischen Verarbeitungen resultierenden Zahlenwerte für die Einschätzungen der Krankheitsbilder legen nahe, in der Weise zu interpretieren, daß nach dem Ausmaß der Angst Krebs sich von den übrigen fünf Krankheitsbildern deutlicher abhebt als diese voneinander. Das ist eine Interpretation, die Intervallskalenqualität voraussetzt. Aus der von den Beurteilern durchgeführten Operation läßt sich eine Intervallskalenqualität jedoch nicht begründen, und es fehlt auch jeder empirische Beleg dafür.

Der Kern der meßtheoretischen Postulate, die am Beispiel der Untersuchung von LEVINE (1962) erläutert worden sind, ist die Entsprechung von numerischem und empirischem Relativ. Das numerische Relativ muß das empirische Relativ exakt repräsentieren. Deswegen wird sie auch von GIGERENZER (1981) "Repräsentationstheorie" genannt, aber als noch immer unzureichender Ansatz kritisiert. GIGERENZER (1981) formuliert diese Kritik so: Messung ist "primär nicht Zuordnung von Zahlen zu empirischen Objekten gemäß einer Vorschrift" ... "sondern psychologische Modellbildung mit Hilfe numerischer Systeme" (S. 11).

Es ist nicht möglich, in allen Implikationen auf diesen über die Meßtheorie im engeren Sinne hinausgehenden modelltheoretischen Ansatz einzugehen. Es soll nur versucht werden, ihn kurz zu resümieren und entscheidende kritische Fragen auch am hier vorgelegten Untersuchungsansatz zu reflektieren, welche sich aus der sogenannten "Modelltheorie der Messung" herauslesen lassen.

Wichtig erscheint einmal der erkenntnistheoretische Grundansatz. Einfacher gesagt beinhaltet er, daß jede Wirklichkeit nicht per se, als sozusagen ontologische Realität erfaßt wird, sondern daß es Menschen nur möglich ist, Modelle der Wirklichkeit zu haben. Dabei "gibt" es schon nicht einmal die Wahrnehmung der Wirk-

lichkeit als solche, sondern nur Wahrnehmungsmodelle der Wirklichkeit (von GIGERENZER expliziert am Beispiel des Phi-Phänomens).

Entsprechend gibt es (nur) Denkmodelle und ebenso emotionale Modellbildungen und so fort. Diese inneren Modellbildungen des einzelnen Individuums (semantische Modelle der ersten Stufe) werden kommunizierbar durch äußere Modellbildungen. Das sind vorzugsweise Modelle im Medium der Sprache, aber auch in anderen Medien (semantische Modelle der zweiten Stufe). Da man über diese Modelle wiederum kommunizieren kann, gibt es weiterhin semantische Modelle höherer Stufen usf.

Gegenstand der Forschung, explizit der psychologischen Forschung, sind die semantischen Modelle der zweiten Stufe ("E", d. h. "empirisches System"), zum Beispiel die Kommunikation des Probanden eines Experimentes mit dem Untersucher, "dem forschenden Subjekt" ("S"), über seine Wahrnehmungen (das ist der Gegenstandsbereich, "Psi"). Dem empirischen System "E" ordnet das "forschende Subjekt S" ein numerisches System "N" zu. Erfüllt sein - empirisch prüfbar erfüllt - muß, daß die Gesetze des numerischen Systems auch für das empirische System gelten. Das entspricht genau der Repräsentationsforderung der klassischen Meßtheorie. Zugleich aber ist zu sehen, daß die Modelltheorie (wie GIGERENZER sie formuliert) über die Repräsentationstheorie hinausgeht, indem sie mehr Elemente in die Betrachtung einbezieht. Zu den vier bereits genannten Elementen kommt noch ein fünftes hinzu: Die Zielsetzung ("Z") des Forschers, welche den Forschungsansatz modelliert.

Die Schlußfolgerung ist: Absolut unzureichend sei, nur die "zweistellige" Modellrelation N : E zu betrachten, wie es die Repräsentationstheorie in Analogie zur klassischen Physik tue. Der psychologischen Forschung sei nur eine fünfstellige Modellrelation angemessen: Ein Subjekt (S) mit einer Zielsetzung (Z) modelliert ein

numerisches System (N) zur Erfassung eines empirischen Systems (E) des Gegenstandsbereiches (Psi).

Insofern ist bereits Messung an sich Modellbildung, und von daher wird die Aussage gefolgert, daß der Gegenstand der Forschung nicht unabhängig vom Modell ist, sondern daß sich Modellbildung durch S und Gegenstand gegenseitig beeinflussen.

GIGERENZER (1981) sieht diese Explikation der Messung als Modellbildung aber noch nicht als hinreichend an, um das Spezifikum psychologischer Forschung zu erfassen. Ein zusätzlicher, noch entscheidenderer Unterschied zum Forschungsansatz der klassischen Naturwissenschaften bestehe in der Eigenart des Forschungsgegenstandes, d. h. des Gegenstandes der Messung. Kritisiert wird, daß klassische psychologische Meßverfahren versuchen, in völliger Analogie zu Naturwissenschaften, etwa der klassischen Physik, zu verfahren. In der Psychologie gelte nicht, daß es Objekte gibt, welche Träger von Eigenschaften sind, etwa in der Art, wie ein Körper im physikalischen Sinne ein Gewicht hat. GIGERENZER (1981) erläutert das an einem Beispiel: Er verweist auf den "trait-Ansatz" in der Persönlichkeitstheorie. Dort werden die Skalenwerte als Ausprägungen der Personen in bestimmten Persönlichkeitsdimensionen interpretiert (Personen = Objekte, Persönlichkeitsdimension = Eigenschaft oder synonym: Merkmal, Relation auf den Objekten). Das sei eine "zweigliedrige Diagnose" (dieser Terminus wohl gemeint im Sinne von "Definition" bzw. "Identifizierung") des Gegenstandes der Messung. Tatsächlich sei aber der Gegenstand der Messung "das urteilende Individuum im Urteil über sich selbst (als Objekt) im Raster vorgegebener Merkmale. Es wird nicht die Persönlichkeit an sich, sondern das Bild eines Individuums von der eigenen Persönlichkeit erfaßt". Die Folgerung ist, daß immer die "triadische Grundeinheit" Individuum x Objekt x Merkmal zu betrachten sei (GIGERENZER S. 94 ff.). Diese Folgerung, gene-

ralisiert auf Skalierungsansätze, bedeutet, daß von den drei Methoden der subjektzentrierten, der reizzentrierten und der reaktionszentrierten Skalierung vor allem die reizzentrierte (= Stimulus-Scaling) sozusagen absolut unpsychologisch und unhaltbar ist.

Insgesamt sind in diesen meßtheoretischen bzw. modelltheoretischen Betrachtungen drei kritische Punkte bzw. Forderungen resümiert worden, die im Hinblick auf die konkret vorliegende Untersuchung zu reflektieren sind.

Die erste kritische Frage folgt aus den Forderungen der Repräsentationstheorie (welche nach GIGERENZER nur die Schwäche hat, daß sie sich sozusagen noch in der babylonischen Gefangenschaft der "zweigliedrigen Diagnose" befindet). Konkret bedeutet die Forderung nach exakter Repräsentation des empirischen Relativs durch das numerische Relativ, daß die statistische Analyse und somit die Interpretation der gewonnenen Daten sich tatsächlich jeweils nur auf dem Skalenniveau bewegen darf, das empirisch belegbar "erreicht" ist. Das ist eine unbestreitbare, klassische Forderung psychologischer Meßmethodik, welche nicht weiter zu diskutieren ist. Die Konsequenz für den methodischen Ansatz dieser hier vorgelegten Untersuchung war: Da die Fragestellung wesentlich auf Interpretation der Abstände von psychologischen Objekten abzielt, ist besser eine sozusagen umständliche und damit zeitraubendere Meßoperation von den Beurteilern zu verlangen. Das läuft auf eine "indirekte" Methode der Skalierung durch die Beurteiler hinaus. Daß etwa die Transitivitätsforderung erfüllt ist, soll empirisch prüfbar bleiben (im Zusammenhang damit auch die sehr wesentliche Prüfung auf Dimensionalität : eindimensional - mehrdimensional), die Intervallskalenqualität soll belegbar sein, und nicht nur - wie in der als Beispiel angeführten Untersuchung von LEVINE (1962) - mit vielleicht noch so guten Gründen nur angenommen werden können. Das hat konkret immer die Konsequenz, daß auf sicher interessante, aber bei be-

grenztem Zeitrahmen dann nicht mehr mögliche Erhebung weiterer Daten (z. B. Erfassung des Intelligenzniveaus der Beurteiler, von Persönlichkeitseigenschaften etc.) zu verzichten ist.

Schwieriger erscheint die auf einen konkreten Untersuchungsansatz bezogene Diskussion der Forderungen, welche sich aus den modelltheoretischen Überlegungen von GIGERENZER (1981) ergeben. Diese Schwierigkeit ist darin begründet, daß die Kritik an bisherigen psychologischen Methodologien an zwei Punkten ansetzt, einmal bezeichnet als das Problem der Modellrelation, zum anderen als Diagnoseproblem. Es scheint aber erstens (ohne daß bei GIGERENZER das explizit zu finden ist), daß das Diagnoseproblem eine nähere Explikation dessen ist, was in der Beschreibung der letzten zwei Stellen der Modellrelation (empirisches System E und Gegenstandsbereich Psi) bereits gesagt ist. Zum zweiten sind die Folgerungen, die aus dem Postulat einer fünfstelligen Modellrelation sich für die Methodik einer Untersuchung ergeben, weniger konkret zu formulieren. Sie überschneiden sich wohl auch teilweise mit denjenigen, die sich aus dem Postulat der Unabdingbarkeit einer dreigliedrigen Diagnose des Gegenstands einer Untersuchung ergeben.

Es soll auf die Theorie der Modellrelation hier nur sehr global eingegangen werden. Sie berührt grundlegende Probleme der Wissenschaftstheorie. Daß GIGERENZER den häufigen Gebrauch von Beispielen oder Analogien der klassischen Physik zur Erläuterung von psychologischen Meß- oder Skaliermethoden einerseits heftig kritisiert, andererseits aber selbst eine wesentliche Erkenntnis der Quanten-Physik, die Unschärfenrelation nach HEISENBERG, als Beleg für seine Thesen anführt, ist widersprüchlich. Denn aus unterschiedlichen Forschungsgegenständen (Naturwissenschaft vs. Sozialwissenschaft - wobei sowohl diese Gegensatzbildung wie auch die damit implizierte Zuordnung der Psychologie bestreitbar ist -) postuliert er einen absolut anderen

Methodenkanon. Es soll zu diesem Widerspruch nur angemerkt sein, daß sich die Quanten-Physik aus der klassischen Physik entwickelt hat, daß weiterhin analog zum Problem der Beeinflussung des konkreten Untersuchungsobjektes durch den Meßvorgang in der Atomphysik in der Psychologie auch im Rahmen der sogenannten klassischen Methodologien Überlegungen und empirische Untersuchungen etwa zum Versuchsleitereffekt, allgemeiner zur Determination des Ergebnisses durch den Untersuchungsansatz, aber auch zum Problem des Meßfehlers (der Zufalls- oder systematischer Fehler sein kann) angestellt worden sind. Wesentlich ist die Überlegung, daß Untersuchungen im Bereich der Psychologie anfällig dafür sind, Artefakte zu produzieren. Das läßt sich aber generell nur durch zusätzliche - konkurrierende - Untersuchungen kontrollieren, kaum innerhalb einer konkreten Untersuchung.

Wiederum auf eine konkrete Untersuchung hin zu bedenken- und von ähnlich entscheidender Bedeutung für die Interpretation von Daten-ist die Frage nach einer Legitimität einer reizzentrierten Skalierung. Es ist bereits erläutert worden, wie GIGERENZER (1981) zu der Schlußfolgerung kommt, eine reizzentrierte Skalierung sei jedem Forschungsgegenstand der Psychologie absolut unangemessen, was im gleichen Zug auch bedeutet, daß die aus einer solchen Methodik gewonnenen Daten nicht sinnvoll interpretierbar sind.

Der Fehler in den Überlegungen GIGERENZERs liegt nicht in der Betonung des spezifischen Forschungsgegenstandes der Psychologie. Auch kann man die Formulierung des sogenannten "Diagnoseproblems" akzeptieren. Falsch ist aber die Ansicht, daß hieraus zwingend folgt, daß in einem einzelnen methodischen Schritt bei der Analyse der Zusammenhänge zwischen den Gliedern Individuum (als Beurteiler oder "Meßinstrument") - Objekt (als Merkmalsträger) - Merkmal (im Sinne von Ausgeprägtheit einer

Eigenschaft an einem Objekt) nicht jeweils wechselseitig einmal die Individuen, ein anderes Mal die Merkmalsausprägung als Quellen der Variation ausgeblendet werden dürften. Die Kritik GIGERENZERs trifft solange nicht, wie nicht mit der methodischen Ausblendung eine prinzipielle Ausblendung von Individuen, also eine prinzipielle Reduktion eines Beurteilers etwa auf die Funktion eines "Meßinstrumentes" oder eine prinzipielle Leugnung der Bedeutsamkeit interindividueller Unterschiede gemeint ist. Nur scheint es, daß GIGERENZER in seinen theoretischen Überlegungen das Bild eines Forschungsansatzes angreift, das er selber entworfen oder zumindest überzeichnet hat. In seinen eingehenden und aufschlußreichen Einzelanalysen von Meßmethoden findet sich eine solche prinzipielle Ablehnung einer Methode auch nicht mehr wieder.

Angemessener wird u. E. das Problem von SIXTL (1972) aufgezeigt. Er kritisiert eine allzu schnelle Ausblendung interindividueller Unterschiede durch Mittelwertsbildung gerade bei reizzentrierten Skalierungsmethoden. Er versucht, empirische Kriterien für die Berechtigung einer solchen methodischen Ausblendung aufzuzeigen (gerade diese Intention verkehrt GIGERENZER (1981) in seiner Zitation dieses Aufsatzes in ihr Gegenteil). So hebt SIXTL (1972) sehr viel anschaulicher den spezifischen Forschungsgegenstand des Psychologen hervor, wenn er als Beispiel einer reizzentrierten Skalierung anführt, daß bei einer Untersuchung etwa zur Skalierung der Helligkeit von Farben der Psychologe sich gar nicht wie der Physiker für die Helligkeit an sich, sondern für den Organismus, welcher Helligkeit beurteilt, interessiert.

Zu beachten sein wird, daß bei allem Interesse an der durchschnittlichen Beurteilung von Krankheitsbildern in der hier vorgelegten Untersuchung so verfahren wird, daß nach SIXTL (1972) potentiell definitive Unterschie-

de der Beurteilung so gut als möglich schon im Untersuchungsansatz abgegrenzt werden von zufälligen Unterschieden. Potentiell definitive Unterschiede könnten etwa zurückgehen auf verschieden gute Kenntnis der Objekte, auf unterschiedliche Bildung bzw. Intelligenz (zu erwarten ist möglicherweise inkonsistentes Beurteilerverhalten), unterschiedliche persönliche Erfahrung mit Krankheiten (im Sinne eher affektiver als kognitiver Prägung einer Einstellung).

3.1.3 Das Prinzip der Methode des Paarvergleichs

Obwohl die Methode der "paired comparisons" ein wenig aus dem Blickfeld geraten ist, erscheint sie als eine der am besten ausgearbeiteten Skalierungsmethoden, was durch die Fülle von methodisch orientierten Arbeiten bis hin zu einer eigenen kleinen Monographie (DAVID 1963) belegt wird. Eine Darstellung der Methode einschließlich der meisten Berechnungsverfahren findet sich in allen klassischen Handbüchern über Skalenkonstruktion bzw. Meßmethoden (KENDALL 1948, GUILFORD 1954, EDWARDS 1957, TORGERSON 1962, SIXTL 1967; 1982), eine geraffte Darstellung der prinzipiellen Schritte bei SÜLLWOLD (1969a).

Beim Paarvergleich wird nicht, wie etwa bei anderen Skalierungsversuchen, einem Beurteiler ein Stimulus (hier: ein Krankheitsbild) gleichzeitig mit einer numerischen oder graphischen Skala dargeboten, oder etwa alle Stimuli gleichzeitig, die dann in eine Rangordnung zu bringen sind, sondern er erhält jeweils ein Paar von Stimuli (jeweils ein Paar von Krankheiten) vorgelegt und soll nur entscheiden (etwa durch Ankreuzen), welcher von beiden Stimuli bzw. welches von beiden Objekten in Bezug auf das genannte Kriterium die stärkere Ausprägung zeigt (hier: welche von beiden Krankheiten er für sich als schlimmer empfindet). Auf diese Weise werden alle möglichen Kombinationen durchgegangen: jeder Stimulus wird mit jedem anderen je

einmal zusammen dargeboten. Dies Verfahren gestattet bereits für den Fall nur eines Beurteilungsdurchganges die Bildung einer Rangordnung der Reize, indem man für jeden Reiz die Zahl der Bevorzugungen addiert. Die klassische Berechnungsmethode sieht dabei vor, den Urteiler insoweit einzuschränken, als ihm das Urteil "gleich" nicht gestattet wird. Weiterentwicklungen gestatten auch Gleichheitsurteile (etwa KAISER & SERLIN 1978).

Läßt man nun (theoretisch) den Beurteiler den Beurteilungsvorgang vielfach wiederholen, so erhält man für jedes Paar die Häufigkeiten der "Bevorzugung" sowohl des einen wie des anderen Paarlings. Die Logik des Verfahrens besteht in folgender Überlegung: Wenn sich die Paarlinge in Bezug auf die ihnen zuzumessende Eigenschaft (das Kriterium, die psychologische Dimension) eindeutig unterscheiden, wird sich eine ganz klare Verteilung der Häufigkeit der Bevorzugung ergeben. So wird der eine Paarling immer (in relativer Häufigkeit zu 1oo % der Beurteilungsfälle) bevorzugt, der andere nie. Je weniger deutlich der Unterschied wird, desto mehr verändert sich diese Verteilung der Bevorzugung: Die Position in Bezug auf die vorgestellte Dimension wird immer häufiger "verwechselt". Für den Fall, daß die gefragte Eigenschaft (Dimension) bei beiden Paarlingen gleich ausgeprägt ist, ist zu erwarten, daß beide mit gleicher Häufigkeit bevorzugt werden. Sollte weiterhin der andere Paarling die stärkere Eigenschaftsausprägung bekommen, so wird die Wahrscheinlichkeit steigen, daß nun er häufiger bevorzugt wird.

Es wurde bereits gesagt, daß man für einen Beurteilungsdurchgang und für einen Beurteiler bereits eine Rangskala bestimmen kann. Bei vielfacher Wiederholung des Beurteilungsvorganges läßt sich eine Intervallskala konstruieren (SIXTL 1967, GIGERENZER 1981), und

zwar dann, wenn man die relativen Häufigkeiten über die Beziehung zwischen relativer Häufigkeit und z-Wert der Standardnormalverteilung (Flächenanteil unter der Kurve und Abszissenmaß) in ebensolche z-Werte umrechnet.

Zu beachten ist, daß man diese Beziehung zwischen relativer Häufigkeit und z-Wert postuliert, indem man eine Urteilsstreuung annimmt, die sich normal verteilt. Das kann man jedoch mit guten inhaltlichen Gründen (GIGERENZER 1981) tun, und ferner muß die Anwendung des Verfahrens durch nachträgliche empirische Prüfung als hinreichend berechtigt nachgewiesen werden. Dieses Postulat ist von THURSTONE (1927) erstmals formuliert worden im "Law of comparative judgement" - Gesetz des Vergleichsurteils. Man sollte bei der Terminologie "Gesetz" bleiben, weil sie inzwischen Standard ist, sollte aber im Bewußtsein behalten, daß es sich um ein Postulat bzw. eine Hypothese handelt, die jeweils zu prüfen und bei Falsifikation zu verwerfen ist.

Die Erklärung ist jetzt noch zu ergänzen: Jeder Reiz, jedes Objekt, jeder Vergleichsgegenstand wird ja nicht nur mit einem, sondern mit allen anderen Objekten verglichen. Man erhält also so viel relative Häufigkeiten und damit z-Werte für ein Objekt, wie es noch andere Vergleichsobjekte gibt. Die Position des Objektes auf der psychologischen Dimension ist aber sinnvoll nur durch einen einzigen Wert zu repräsentieren. Dafür bietet sich das arithmetische Mittel der z-Werte für ein Objekt an. Das arithmetische Mittel ist der Wert, der deswegen der "repräsentativste" ist, weil die Summe der Abweichungen der Einzelwerte von diesem Mittelwert null ist bzw. die Summe der Abweichungsquadrate die kleinstmögliche (ein Minimum) (SIXTL 1967).

Bisher ist eher der elementare Berechnungsvorgang betrachtet als die vollständige logisch mathematische Ableitung (dazu vgl. u. a. SIXTL 1967). Nun ist noch aus-

drücklich zu betonen, daß der z-Wert ja die Umsetzung eines relativen Häufigkeitswertes (p) ist, dieser Häufigkeitswert aber psychologisch eine Differenzbildung zwischen zwei Reizen, also einen Diskriminationsprozeß voraussetzt. Ganz zu Anfang dieser Erläuterung wurde davon gesprochen, daß die Verteilung der relativen Häufigkeiten umso extremer ist (sich von o.5 zu o.5 immer mehr in Richtung 1.o zu o.o hin verschiebt), je deutlicher die Reize unterschieden sind. Auf diesem Hintergrund wird die einfachste mathematische Formulierung des "Law of comparative judgement" von THURSTONE (1927) verständlich:

$$R_j - R_k = z_{jk} \cdot$$

Das bedeutet: Der z-Wert (z_{jk}) repräsentiert den Abstand (-) der beiden Objekte (R_j; R_k) auf der psychologischen Dimension. Durch die oben beschriebene Mittelwertsbildung der z-Werte für ein Objekt läßt sich dann die Position des Objektes auf der Skala (die durch die Relation aller Objekte zueinander gebildet wird!) schätzen.

Die oben zitierte Formel (GUILFORD 1954) ist die einfachste Form des Law of comparative judgement, der sogenannte Fall V. In seiner vollständigen Form (Fall I und II) ist es so formuliert:

$$R_j - R_k = z_{jk} \cdot \sqrt{\sigma_j^2 + \sigma_k^2 - 2r_{jk}\,\sigma_j\,\sigma_k} \cdot$$

Beim Vergleich beider Formeln sieht man, daß man für den Fall V einige zusätzliche Annahmen machen muß, um die Größen, welche die Varianzen bzw. Standardabweichungen des Diskriminationsprozesses und die mögliche Korrelation bei der Einstufung von Objekten bezeichnen, fortlassen zu können. Wir hatten davon gesprochen, daß psychologisch gesehen eine Einstufung eines Reizes auf einer Skala, wenn sie vielfach vorgenommen ist, sich

verteilt, d. h. um einen Wert streut, woraus folgt, daß Verteilungen sich überlappen können - sonst gäbe es nur relative Häufigkeiten 1.o zu o.o, womit die Anwendung des Paarvergleichs unmöglich wäre.

Die erste zusätzliche Annahme, nämlich daß die Korrelation außer Betracht bleiben kann, weil sie null ist, weil also kein Zusammenhang bei der Einstufung von je zwei Objekten besteht, führt zu einer ersten Vereinfachung, zu Fall III:

$$R_j - R_k = z_{jk} \cdot \sqrt{\sigma_j^2 + \sigma_k^2}$$

(Diese Formel ergibt sich, wenn in die Formel für Fall I/II $r_{jk} = o$ gesetzt wird.)

Auf den zuerst erwähnten Fall V, die einfachste Form, führt die weitere zusätzliche Annahme, daß die Varianzen der Urteilsstreuungen gleich sind. Der Ausdruck unter der Wurzel verändert sich dann zu $2\,\sigma_j^2$, und wenn man, was der Standardnormalverteilung entspricht $2\,\sigma_j^2 = \sigma_j \cdot \sqrt{2}$ als Einheit nimmt, also gleich 1 setzt, erhalten wir die Formel für Fall V.

Der Vollständigkeit halber soll noch angeführt werden, was es genauer mit Fall I und Fall II und mit Fall IV auf sich hat.

Fall I und II implizieren die vollständige Anwendung der Formel ohne Vorannahmen, welche sich aber nicht als praktikabel erwiesen hat, da man keinen empirischen Ansatz finden kann r_{jk} zu schätzen. Im Unterschied zu Fall I, der das Modell - wie wir es anfangs beschrieben haben - in einer vielfachen Wiederholung bei einem Beurteiler beinhaltet, bezieht sich Fall II auf eine Beurteilung durch eine Vielzahl von Beurteilern (wie in unserer Untersuchung durchgeführt). Fall IV ist ähnlich Fall V, nur daß nicht Gleichheit, sondern nur annähernde Gleichheit der Varianzen angenommen wird (GUILFORD 1954).

Auch dieser Fall ist nicht von praktischem Interesse.

Wie zu sehen ist, unterscheiden sich die Fälle und damit die Lösungsmethoden nach der Art und Anzahl der Vorannahmen über die Verteilungscharakteristika der "diskriminativen Prozesse", die dem Paarvergleichs-Urteil zugrunde liegen. Es sei noch erwähnt, daß andere Autoren (GOUDE 1962 und EISLER 1965, zitiert nach GIGERENZER 1981) einen Fall VI vorschlagen, bei dem eine logarithmische Normalverteilung für die Diskriminationsprozesse angenommen wird.

Es ergeben sich grundsätzliche Fragen nach der Berechtigung der geschilderten Vorannahmen. Zur Berechtigung ist zu sagen, daß sie aus den Daten heraus exakt geprüft wird. Beim Rechenverfahren, das in seinen logischen Grundzügen bereits geschildert ist (eine detaillierte Ableitung und Darstellung ist zu finden bei SIXTL 1967; Rechenverfahren bei GUILFORD 1954; KENDALL 1948; TORGERSON 1962; konkrete meßtheoretische Untersuchung bei GIGERENZER 1981), werden die empirischen Häufigkeiten in einer quadratischen Matrix angeordnet, und zwar in der Art, daß sowohl die Zeilen wie die Spalten durch die zu vergleichenden Objekte bezeichnet werden. In die dadurch entstehenden Zellen werden dann die absoluten Häufigkeiten ("Häufigkeitsmatrix") der Bevorzugung des die Spalte bezeichnenden Objekts über das die Zeile bezeichnende Objekt eingetragen. Aus dieser empirischen Häufigkeitsmatrix wird die Matrix der relativen Häufigkeit entwickelt, daraus die Matrix der z-Werte und daraus (mit oder ohne Einbeziehung einer Varianzschätzung) die Skala der Objekte. Zur empirischen Prüfung wird im Prinzip dann von der Skala der Objekte eine theoretische Matrix zurückgerechnet, "reproduziert" (SIXTL 1967). Diese kann mittels Signifikanztest mit der empirischen Ausgangsmatrix verglichen werden, wodurch die Berechtigung des Verfahrens insge-

samt (Eindimensionalität) wie auch zusätzliche Annahmen geprüft werden können.

3.1.4 Rechenverfahren und Prüfung der Skalierungshypothesen am Beispiel der Gruppe der Medizinstudenten (Med 1)

3.1.4.1 Prüfung der individuellen Urteilskonsistenz

Aus den bisherigen Ausführungen ergibt sich, daß gute Gründe dafür vorhanden sind, für unsere Frage nach einer Skala der subjektiven Einschätzung des Schweregrades ("Schlimmheit") von Krankheiten, insbesondere von psychiatrischen Erkrankungen, die Paarvergleichstechnik auszuwählen. Die Gründe sind aber logische Gründe. Daß Einstellungsskalen zu verschiedenen Völkern (SÜLLWOLD 1967) mit diesem Verfahren erfolgreich zu erstellen sind, ist für uns ein methodisches Vorbild gewesen. Es beinhaltet aber noch nicht, daß das Modell auf unser Problem wirklich paßt. Die Grundfrage ist, ob auch die "Schlimmheit" von Krankheiten eindimensional skalierbar ist. Mehr praktisch sind die Fragen, ob die vorgegebene Instruktion, welche die psychologische Dimension als Kriterium einführt, adäquates Beurteilerverhalten provoziert.

Eine erste empirische Prüfung, ursprünglich als Voruntersuchung gedacht, nahmen wir in der Weise vor, daß wir den Medizinstudenten der Medizinischen Fakultät des Saarlandes, die im Sommersemester 1980 den Kursus in Psychiatrie begannen, die Liste der Paar-Vergleiche vorlegten. Anhand dieser empirischen Daten zu folgenden somatischen Krankheitsbildern:

Grippe,
Akne,
Blinddarmentzündung,
Bluthochdruck,
Zuckerkrankheit,

Syphilis,
Herzinfarkt,
Tuberkulose,
Querschnittslähmung,
Krebs,

und folgenden "untergemischten" psychiatrischen Krankheitsbildern:

Neurose,
Alkoholismus,
endogene Depression,
Hirnarteriosklerose,
Schizophrenie,

erhielten wir von N = 66 Medizinstudenten (im folgenden als Gruppe Med 1 bezeichnet) den ersten empirischen Datensatz. Für die Auswertung auf den Rechnern TR 44o und Siemens 751 des Rechenzentrums der Universität des Saarlandes benutzten wir ein Fortran IV-Programm für das Verfahren nach Fall V und Fall III der GMD, wobei die Form der Dateneingabe von uns modifiziert wurde, eine Ergänzung der Prüfverfahren vorgenommen wurde, sowie nach der von SIXTL (1967) beschriebenen Rechenvorschrift die Berechnung der Skala nach dem Iterationsverfahren von GULLIKSEN (1956) angefügt wurde *).

---

*)

Zu danken ist der Gesellschaft für Mathematik und Datenverarbeitung, früher Deutsches Rechenzentrum, Darmstadt, für die freundliche Überlassung des Programms von M. KOLBE. Ganz besonders zu bedanken haben wir uns jedoch bei Herrn Dipl.-Math. S. GRÄBER, Nebenstelle Homburg des Rechenzentrums der Universität d. Saarlandes, für seine fachliche Beratung sowie für die Erstellung eines Fortran IV-Programms für das Iterationsverfahren und die Ergänzungen und Modifikationen des zugrunde gelegten Programms nach unseren Wünschen. Ohne seine hervorragende Hilfsbereitschaft wäre die jetzt komfortable Verarbeitung der Daten nicht möglich gewesen.

Die weiteren statistischen Berechnungen wurden mit den entsprechenden Prozeduren des Programm-Paketes SPSS (NIE, HULL, JENKINS, STEINBRENNER & BENT 1975; BEUTEL, KÜFFNER & SCHUBÖ 198o) durchgeführt.

Die Methode des Paar-Vergleichs gestattet, wie bereits erläutert, eine Prüfung der Kompetenz des einzelnen Beurteilers, indem für jeden Beurteiler die Anzahl sogenannter "zirkulärer Triaden" (KENDALL 1955, SÜLLWOLD 1969a,SIXTL 1967) bestimmt wird. Unter einer "zirkulären Triade" versteht man einen logischen Fehler der Art, daß die Transitivitätsforderung (GIGERENZER 1981) verletzt wird. Beim Vergleich der Krankheiten ist das der Fall, wenn etwa ein Beurteiler Neurose als schlimmer ansieht als Bluthochdruck und weiterhin Zuckerkrankheit als schlimmer beurteilt als Neurose, dann aber nicht, wie jetzt logisch zwingend, Zuckerkrankheit als schlimmer als Bluthochdruck, sondern umgekehrt Bluthochdruck schlimmer als Zuckerkrankheit. Formalisiert dargestellt:

Wenn: Neurose > Bluthochdruck
und : Neurose < Zuckerkrankheit
(äquivalent: Zuckerkrankheit > Neurose)

ergibt das die Reihe
Zuckerkrankheit > Neurose > Bluthochdruck;
es folgt also:
Zuckerkrankheit > Bluthochdruck.

Das Urteil "Bluthochdruck > Zuckerkrankheit" macht aber das Dreierpaar zu einer zirkulären Triade, zu einer intransitiven Urteilsreihe. SIXTL (1967) und SÜLLWOLD (1967; 1969a) fassen die Gründe, die für das Auftreten von zirkulären Triaden gefunden worden sind, zusammen. Danach sind folgende aufzuführen, wobei die ersten drei von besonderem Interesse sind:

1. Inkompetenz des Beurteilers aufgrund mangelnden

Verständnisses, großer Unklarheit über die Objekte oder Mangel an Sorgfalt.

2. Wenn die Objekte nicht in einer Dimension zu skalieren sind, d. h. wenn naheliegt, das Kriterium der Beurteilung ("Schlimmheit") bei verschiedenen Paaren anders inhaltlich zu füllen (bei unserem Problem etwa: schlimm im Hinblick auf physische Gefährdung, schlimm im Hinblick auf soziale Folgen). Veranschaulichen läßt sich das mit einem Beispiel: von drei Bergen kann je nach Standpunkt bzw. Blickrichtung des Beobachters (ohne rationale Korrektur nach den Gesetzmäßigkeiten des perspektivischen Sehens) mal der eine, mal der andere als höchster Berg erscheinen. Eine generell hohe Zahl von zirkulären Triaden kann also auf mangelnde Eindimensionalität hinweisen.

3. Zirkuläre Triaden treten gehäuft auf, wenn die Differenz zwischen Objekten auf der Skala sehr klein ist, wenn es also wahrscheinlicher ist, daß Objekte in ihrer Skalenposition "verwechselt" werden.

4. Schließlich gibt es Überlegungen und auch empirische Hinweise darauf, daß eine Häufung von zirkulären Triaden persönlichkeitsspezifisch sein könnte (HILL 1953, vgl. SIXTL 1967; SÜLLWOLD 1969a). Hierzu ist in Nähe zu 1., aber in einer anderen, klinischen Akzentuierung zu erwähnen, daß BRAATZ (197o) die Häufigkeit von intransitiven Urteilen als Indikator für Denkstörungen bei Schizophrenie vorschlägt.

KENDALL (1955) hat eine Methode entwickelt, die bei dem einzelnen Beurteiler eines Paarvergleichs tatsächlich vorkommende Zahl von zirkulären Triaden zu berechnen. Dazu ist für jeden Beurteiler die sog. "Dominanzmatrix" zu erstellen. Hierbei werden in einer quadratischen Matrix genau wie bei der bereits beschrie-

benen Häufigkeitsmatrix für die gesamte Untersuchungsgruppe (welche die Summe aller einzelnen Dominanzmatrixen ist) Spalten und Zeilen mit den Objekten in gleicher Reihenfolge gekennzeichnet. Wenn dann der Beurteiler das Spaltenobjekt als die stärkere Ausprägung des Kriteriums repräsentierend beurteilt, wird in die entsprechende Zeile eine 1 eingetragen, sonst eine o. Die Spaltensummen (a_j) ergeben die Rangordnung, die der Beurteiler den Objekten gibt. Aus diesen Spaltensummen (a_j) ist nach der KENDALLschen Formel die Anzahl (d) der "zirkulären Triaden" zu bestimmen:

$$d = (n \cdot (n-1) \cdot (2n-1)/12) - 1/2 \cdot \sum_{j=1}^{n} a_j^2 .$$

Während das erste Glied der Formel nur durch die Anzahl der Objekte n bestimmt ist, hängt die numerische Größe des zweiten von der Anzahl gleicher Rangplätze ab: Je mehr gleiche Rangplätze (wobei der niedrigste Rangplatz hier mit o zu zählen ist), desto kleiner wird dieses zweite Glied, d. h. wenn der Beurteiler keine Systematik seines Urteilens zeigt, wenn er sozusagen keine klare Vorstellung von einer Rangordnung ohne gleiche Rangplätze hat (etwa aus einem oder mehreren der oben genannten Gründe), wird er intransitive Urteile fällen.

Ein numerisches Beispiel: Bei fünf Objekten mit den Rangplätzen 1, 2, 3, 4 und 5, ergibt sich

$$\begin{aligned} d &= 5 \times 4 \times 9/12 - 1/2 \ (o^2 + 1^2 + 2^2 + 3^2 + 4^2) \\ &= \quad 15 \quad - 15 \\ &= o \end{aligned}$$

Weiter schlägt KENDALL (1948) eine Formel für die Berechnung der maximal möglichen Anzahl von zirkulären Triaden vor:

bei gerader Anzahl von Objekten: $d_{max} = n\,(n^2 - 4)/24$
bei ungerader Anzahl von Objekten: $d_{max} = n\,(n^2 - 1)/24$

Nach KENDALL (1948) ist dann aus dem Verhältnis der Anzahl tatsächlicher zirkulärer Triaden zur Anzahl der maximal möglichen ein "Koeffizient der Konsistenz" des Beurteilungsverhaltens zu bestimmen, nach der Formel Zeta = 1 - d/d_{max}, bzw. das Ausmaß der Konsistenz nach Chi² zu testen (bei mehr als sechs Objekten; KENDALL 1948, SIXTL 1967, LIENERT 1978). Dabei wird gegen die Null-Hypothese getestet, daß der Grad an Konsistenz in einer Dominanzmatrix nur als Zufallsergebnis zu betrachten sei.

SIXTL (1967) argumentiert hingegen, daß die Bestimmung von d_{max} nicht ganz zutreffend sei, da diese Zahl nicht nur von der Anzahl n der Objekte abhänge, sondern auch davon, wie ein Beurteiler das Objekt, das in den beiden ersten Paaren eines Tripels (Dreiergruppe) vorkommt, einordne. Das dritte Urteil liege logisch nur dann fest, wenn der Beurteiler dieses Objekt der Rangordnung nach in die Mitte zwischen den beiden anderen einordne.

Eine Nachprüfung zeigt, daß SIXTL zu Recht argumentiert. Wenn bei drei Objekten A, B und C ein Paarvergleich durchgeführt wird, ergeben sich folgende möglichen Zweierkombinationen:

A - B
A - C und
B - C.

Falls der Beurteiler die beiden ersten Vergleiche "gleichsinnig" entscheidet, also beide Paare wie folgt beurteilt:

A größer B, A größer C

oder umgekehrt

A kleiner B, A kleiner C,

kann er beim Paar B - C sowohl "größer" als auch "kleiner" urteilen, ohne ein intransitives Urteil zu fällen. Das kann vielmehr nur resultieren, wenn er bei den ersten beiden Paaren "gegensinnig" entscheidet, also

A größer B, A kleiner C

oder umgekehrt

A kleiner B, A größer C.

Nur bei dieser Urteilsanordnung folgt logisch, daß dann C entweder "größer B" sein muß, oder im umgekehrten Fall "kleiner B". Nur hier wird A zwischen die beiden anderen Objekte eingeordnet. Solch eine Einordnung ist nach SIXTL (1967) aber nur in der Hälfte aller möglichen Anordnungen gegeben. Daraus folgt, daß die Zahl der maximal möglichen zirkulären Triaden jeweils pro Beurteiler einzeln zu bestimmen wäre.

Wir haben daraus die Konsequenz gezogen, das Kriterium für die Urteilskonsistenz schärfer anzusetzen, als es nach KENDALLs Chi² notwendig wäre. Für unsere 15 Krankheitsbilder ergibt sich nach KENDALL d_{max} = 15 $(15^2-1)/24$ = 14o.

Für die Chi²-Prüfung bei einer Irrtumswahrscheinlichkeit von α= .o1 gegen H_o folgt ein Chi² von 41.64, das erreicht werden müßte. Dieses Chi² wird erreicht bei einer Anzahl von 89 zirkulären Triaden (diese Zahl hängt nur von der Anzahl 15 der Objekte ab und ist für alle Beurteiler gleich; sie ergibt sich aus der KENDALLschen Chi²-Formel). Um den Überlegungen von SIXTL Rechnung zu tragen, haben wir deswegen nicht nach Zeta oder Chi² geurteilt, sondern festgelegt, daß alle die Beurteiler ausscheiden, die mehr als die Hälfte von 14o möglichen, also mindestens 71 zirkuläre Triaden erreichen (dieser Wert überschreitet noch das Signifikanzniveau von = .oo1). Bei der Gruppe Med 1 mußte danach kein Beurteiler ausscheiden. Es fand sich immerhin bei vier von 66 Beurteilern keine zirkuläre Triade, 45 hatten 1 bis 1o zirkuläre Triaden, 15 Beurteiler hatten 11 bis 2o, einer 26 und einer 32. Die Medizinstudenten urteilten also hochkonsistent.

3.1.4.2 Die Berechnungsverfahren nach Fall V und Fall III und die Prüfung der Güte der Anpassung

Im eigentlichen Berechnungsverfahren wird nun zuerst die Matrix der Häufigkeiten erstellt, welche für die Stichprobe Med 1 hier vollständig wiedergegeben werden soll (Tab. 1).

Bei einer Inspektion der Daten (Tab. 1) ist zu sehen, daß relativ viele extreme Häufigkeitsverteilungen vorkommen. Das läßt bereits vermuten, daß Lösungen nach Fall V und Fall III keine optimale Anpassung an die empirischen Daten bringen werden. In der Matrix ist weiterhin zu sehen, daß in die Diagonale jeweils n/2 als Häufigkeit eingesetzt wird (was zu einem z-Wert von o.o führt), und daß aufgrund der Instruktion "gleich schlimm ist nicht gestattet" eine Korrespondenz einander zur Diagonalen symmetrischer Zellen erzwungen wird, die summiert immer N (66) ergeben.

Die Matrix der absoluten Häufigkeiten wird dann, wie bereits beschrieben, umgewandelt in die Matrix der relativen Häufigkeiten (p-Matrix) und im nächsten Schritt entsprechend der einfachsten Formulierung des Gesetzes des Paar-Vergleichs in die Matrix der z-Werte umgerechnet (Tab. 2). Auch hier entsprechen sich die Werte in den zueinander symmetrischen Zellen; durch die Umrechnung sind sie numerisch gleich, haben nur entgegengesetztes Vorzeichen. Die z-Werte einer Spalte sind zu interpretieren als Abstandswerte auf der psychologischen Dimension des Spaltenobjekts jeweils von allen anderen (Abstand o.o von sich selbst). Der Skalenwert für Fall V (Tab. 3) ist der Mittelwert aller z-Werte einer Spalte. Er repräsentiert den Abstand einer Krankheit auf der psychologischen Dimension "Schlimmheit" von allen anderen Krankheiten, d. h. also vom Gesamtmittel.

Dieser Sachverhalt ist auch für eine Interpretation

Tabelle 1:

Matrix der absoluten Häufigkeiten aus einem Paarvergleich

(Daten der Gruppe Med 1, N = 66; Zahlen = Anzahl der Beurteiler, die das die Spalte bezeichnende Objekt als "schlimmer" beurteilen als das die Zeile bezeichnende Objekt)

	GRIP	AKNE	NEUR	BLIN	BLUT	ALKO	ZUCK	SYPH	HIRN	TB	MERZ	DEPR	QUER	KRBS	SCHI
GRIP	33,0	40,0	64,0	53,0	63,0	65,0	63,0	64,0	66,0	65,0	66,0	64,0	66,0	66,0	66,0
AKNE	26,0	33,0	66,0	49,0	61,0	65,0	63,0	62,0	66,0	65,0	66,0	66,0	66,0	64,0	65,0
NEUR	2,0	0,0	33,0	10,0	25,0	44,0	27,0	31,0	51,0	40,0	59,0	48,0	61,0	64,0	61,0
BLIN	13,0	17,0	56,0	33,0	58,0	62,0	60,0	61,0	62,0	64,0	66,0	62,0	60,0	65,0	66,0
BLUT	3,0	5,0	41,0	8,0	33,0	47,0	42,0	39,0	60,0	54,0	63,0	55,0	66,0	65,0	60,0
ALKO	1,0	1,0	22,0	4,0	19,0	33,0	16,0	28,0	44,0	25,0	45,0	33,0	58,0	62,0	54,0
ZUCK	3,0	3,0	39,0	6,0	24,0	50,0	33,0	40,0	60,0	43,0	62,0	43,0	64,0	66,0	60,0
SYPH	2,0	4,0	35,0	7,0	27,0	38,0	26,0	33,0	49,0	27,0	55,0	36,0	61,0	63,0	59,0
HIRN	0,0	0,0	15,0	4,0	6,0	22,0	6,0	17,0	33,0	24,0	38,0	26,0	54,0	59,0	41,0
TB	1,0	1,0	26,0	2,0	12,0	41,0	23,0	39,0	42,0	33,0	55,0	40,0	66,0	66,0	51,0
MERZ	0,0	0,0	7,0	0,0	1,0	21,0	4,0	11,0	28,0	11,0	33,0	13,0	53,0	64,0	36,0
DEPR	2,0	0,0	18,0	4,0	11,0	33,0	23,0	30,0	40,0	26,0	53,0	33,0	58,0	61,0	54,0
QUER	0,0	0,0	5,0	0,0	0,0	8,0	2,0	5,0	12,0	0,0	11,0	8,0	33,0	38,0	15,0
KRBS	0,0	2,0	2,0	1,0	1,0	4,0	0,0	3,0	7,0	0,0	2,0	5,0	28,0	33,0	8,0
SCHI	0,0	1,0	5,0	0,0	6,0	12,0	6,0	11,0	25,0	15,0	30,0	12,0	51,0	58,0	33,0

Tabelle 2:

Matrix der z-Werte aus einem Paarvergleich

(Daten der Tab. 1; Wert 0.0 außerhalb der Diagonalen gilt als "nicht beobachtet")

	GRIP	AKNE	NEUR	BLIN	BLUT	ALKO	ZUCK	SYPH	MIRN	TB	HERZ	DEPR	QUER	KNOS	SCHI
GRIP	0.0	0.269	1.876	0.852	1.691	2.166	1.691	1.876	0.0	2.166	0.0	1.876	0.0	0.0	0.0
AKNE	-0.269	0.0	0.0	0.651	1.434	2.166	1.691	1.550	0.0	2.166	0.0	0.0	0.0	1.876	2.166
NEUR	-1.876	0.0	0.0	-1.030	-0.309	0.431	-0.230	-0.076	0.748	0.269	1.248	0.605	1.434	1.876	1.434
BLIN	-0.852	-0.651	1.030	0.0	1.169	1.550	1.335	1.434	1.550	1.876	0.0	1.550	0.0	2.166	0.0
BLUT	-1.691	-1.434	0.309	-1.169	0.0	0.560	0.349	0.230	1.335	0.908	2.166	0.967	0.0	2.166	1.335
ALKO	-2.166	-2.166	-0.431	-1.550	-0.560	0.0	-0.699	-0.191	0.431	-0.309	0.473	0.0	1.169	1.550	0.908
ZUCK	-1.691	-1.691	0.230	-1.335	-0.349	0.699	0.0	0.269	1.335	0.389	1.550	0.389	1.876	0.0	1.335
SYPH	-1.876	-1.550	0.076	-1.434	-0.230	0.191	-0.269	0.0	0.651	-0.230	0.967	0.114	1.434	1.691	0.967
MIRN	0.0	0.0	-0.748	-1.550	-1.335	-0.431	-1.335	-0.651	0.0	-0.349	0.191	-0.269	0.908	1.248	0.309
TB	-2.166	-2.166	-0.269	-1.876	-0.908	0.309	-0.389	0.230	0.349	0.0	0.967	0.269	0.0	0.0	0.748
HERZ	0.0	0.0	-1.248	0.0	-2.166	-0.473	-1.550	-0.967	-0.191	-0.967	0.0	-0.852	0.967	1.876	0.114
DEPR	-1.876	0.0	-0.605	-1.550	-0.967	0.0	-0.389	-0.114	0.269	-0.269	0.852	0.0	1.169	1.434	0.908
QUER	0.0	0.0	-1.434	0.0	0.0	-1.169	-1.876	-1.434	-0.908	0.0	-0.967	-1.169	0.0	0.191	-0.748
KNOS	0.0	-1.876	-1.876	-2.166	-2.166	-1.550	0.0	-1.691	-1.248	0.0	-1.876	-1.434	-0.191	0.0	-1.169
SCHI	0.0	-2.166	-1.434	0.0	-1.335	-0.908	-1.335	-0.967	-0.309	-0.748	-0.114	-0.908	0.748	1.169	0.0

Tabelle 3:

Berechnung der Skalenwerte nach Fall V am Beispiel der Daten der Gruppe Med 1

	Summe z (Spalten-summe Tab.2)	$\bar{z}$ (arithmet. Mittel aus beobachteten Werten *))	Skalenwert (justiert z + 1.446)
Grip	-14.464	-1.446	o.ooo
Akne	-13.431	-1.343	o.1o3
Neur	- 4.524	-o.323	1.123
Blin	-12.157	-1.o13	o.433
Blut	- 6.o32	-o.431	1.o16
Alko	3.54o	o.253	1.699
Zuck	- 3.oo8	-o.215	1.232
Syph	- o.5o3	-o.o34	1.413
Hirn	4.o11	o.3o9	1.755
Tb	4.9o4	o.377	1.824
Herz	5.457	o.455	1.9o1
Depr	1.138	o.o88	1.534
Quer	9.515	o.952	2.398
Krbs	17.244	1.437	2.883
Schi	8.3o9	o.639	2.o86

*) Anm.: Relative Häufigkeiten 1.o zu o.o lassen keine Überschneidung der Objekte mehr beobachten; es wird daher z = o.ooo gesetzt und dieser Wert bei der Mittelung ausgelassen

von Bedeutung: Die Beurteilung eines Reizes geschieht immer in Relation zum gesamten umgebenden Reizfeld. Das bedeutet zugleich, daß Reduktionen der alten Zahl von 15 auf etwa 1o auf rechnerischem Weg nicht ohne weiteres ein identisches Ergebnis erbringen dürften wie eine alleinige Vorgabe von nur 1o Items.

Für den Chi^2-Signifikanztest nach MOSTELLER (1951) muß nun eine "theoretische" Matrix aus den Skalenwerten zurückgerechnet werden. Dabei begnügt man sich mit der Bestimmung der oberen oder unteren Halbmatrix (der durch die Diagonale geteilten quadratischen Matrix), da die Matrix der z-Werte ja den Absolutbeträgen nach symmetrisch ist. Einander entsprechende Werte unterscheiden sich nur durch das Vorzeichen. Den Skalenwert eines Objektes haben wir definiert als Abstand vom Mittelwert aller Objekte, was identisch ist mit dem Mittelwert aus den Abständen dieses einen Objektes von allen anderen. Zugleich gilt, daß der Abstand zweier Objekte voneinander gleich ist der Differenz der Mittelwerte aus den Abständen beider Objekte von allen anderen (SIXTL 1967). Aufgrund dieser Regeln läßt sich aus den Skalenwerten die "theoretische" Matrix der z-Werte reproduzieren, indem die Zellen der Matrix den Abstandswert des Spaltenobjekts vom Zeilenobjekt enthalten. Rechnerisch geschieht das dadurch, daß in die erste Spalte und in die erste Zeile in gleicher Reihenfolge die Skalenwerte eingetragen werden. Durch Subtraktion des Wertes am Kopf der Spalte vom Wert am Anfang der Zeile erhält man den Wert in der entsprechenden Zelle und damit die reproduzierte z-Matrix, die dann in Umkehrung des anfänglichen Rechenverfahrens in die Matrix der relativen Häufigkeiten umgerechnet wird, und zwar wiederum über die Beziehung zwischen z-Skalenwert unter der Standardnormalkurve und zugehörigem Flächenanteil. Eine weitere Transformation sowohl der empirischen als auch der

"theoretischen" Matrix der relativen Häufigkeiten p nach der Formel

$$\text{Theta} = \arcsin \sqrt{p}$$

ermöglicht den Chi²-Test nach MOSTELLER (1951). Er prüft, ob die Abweichungen als zufällig zu beurteilen sind und ergänzt zufallskritisch so den bereits von THURSTONE (1927) praktizierten Vergleich von empirischer und aus den Skalenwerten rekonstruierter Matrix.

Die Berechnung des Chi²-Wertes für die nach Fall V reproduzierte Matrix ergibt bei den Daten der Gruppe Med 1 ein hochsignifikantes Ergebnis (Chi² = 749.o68). Danach ist also eine Anpassung nicht hinreichend erreicht.

Das Rechenverfahren sieht eine erneute Bestimmung der Skalenwerte jetzt nach Fall III unter Einbeziehung der geschätzten (und jetzt nicht mehr als gleich angenommenen) Varianzen der Reaktionsverteilungen vor. Die Ableitung der Formel und die Schätzung der Varianzen aus der Varianz jeweils einer Spalte der Matrix findet sich bei SIXTL (1967). Aus der Skala nach Fall III (Tab. 4) wird nun erneut, diesmal jedoch unter Verwendung der zugehörigen Varianzen, eine reproduzierte "theoretische" z-Matrix und dann p-Matrix errechnet, die wiederum nach MOSTELLER (1951) auf die Güte der Anpassung an die empirische Matrix geprüft werden kann.

Für die Untersuchungsgruppe Med 1 zeigt sich wiederum eine nach dem Chi²-Kriterium von MOSTELLER (1951) hochsignifikante Abweichung (Chi² = 91o.427).

Es sind die Gründe für diese mangelnde Anpassung zu untersuchen. Zur Erklärung bieten sich drei Möglichkeiten an:

1. "Schlimmheit" von Krankheiten ist nicht linear skalierbar, die Methodik ist also dem Problem nicht adäquat.
2. Es gibt zu viele Extremverteilungen; die Skalie-

Tabelle 4:

Gruppe Med 1: Vergleich der Skalenwerte nach den drei Berechnungsverfahren Fall V, Fall III und Iteration

(Krankheiten geordnet nach Skalenposition; Pfeile bezeichnen Wechsel der Rangplätze) und Produkt-Moment-Korrelationen r sowie Spearman-Rangkorrelationen R_S.

Skala V		Skala III		Iteration
Grip o.ooo	———	Grip o.ooo	———	Grip o.ooo
Akne o.1o3	———	Akne o.429	———	Akne o.355
Blin o.433	———	Blin 1.oo3	———	Blin o.7o8
Blut 1.1o6	↘	Neur 1.752	↗	Blut 1.454
Neur 1.123	↗	Blut 1.811	↘	Zuck 1.66o
Zuck 1.232	———	Zuck 1.976	↘	Neur 1.784
Syph 1.413	———	Syph 2.125	———	Syph 1.957
Depr 1.534	———	Depr 2.257	↘	Tb 2.129
Alko 1.699	———	Alko 2.452	↘	Depr 2.186
Hirn 1.755	↘	Herz 2.5o3	↘	Alko 2.23o
Tb 1.824	↗	Hirn 2.577	→	Hirn 2.578
Herz 1.9o1	↘	Tb 2.8o4	↗	Herz 2.855
Schi 2.o86	———	Schi 2.984	———	Schi 2.882
Quer 2.398	———	Quer 4.441	———	Quer 3.473
Krbs 2.883	———	Krbs 5.o51	———	Krbs 3.61o

Produkt-Moment-Korrelationen:

r	III	It
V	.971	.986
III		.953

Spearman-Rangkorrelationen:

R_S	III	It
V	.986	.975
III		.95o

rungsmethoden nach Fall V oder Fall III von THURSTONE sind nicht adäquat.

3. Der Test für die Güte der Anpassung nach MOSTELLER ist nicht adäquat.

Der Grund unter (2) ist mit Sicherheit zutreffend. Schon eine Inspektion der Daten (Tab. 2) hat gezeigt, daß die Beurteilergruppe deutlich bis eindeutig zwischen schweren Krankheiten und leichten Krankheiten differenziert. Die Reaktionsverteilungen überlappen sich nicht mehr. Für ein Verhältnis der relativen Häufigkeiten von 1.o zu o.o gibt es aber keine repräsentierenden z-Werte unter der Standardnormalkurve. Wir haben bisher diesen Fall so behandelt, dann die Mittelung durch n-1 (Anzahl der Items weniger 1; bzw. n-2 usw.) vorzunehmen *).

Als Ergebnis ist bisher festzuhalten, daß weder nach Fall III noch nach Fall V eine adäquate Lösung zu erzielen ist.

3.1.4.3 Das Iterationsverfahren nach GULLIKSEN und die Indices der Skalierbarkeit

Das Lösungsverfahren, das solchen Extremverteilungen gerecht wird, ist das von GULLIKSEN (1956) entwickelte Iterationsverfahren, das bei größeren Itemzahlen einzig den Nachteil hat, daß es erhebliche Rechenarbeit voraussetzt, andererseits den beträchtlichen Vorteil, daß es sogar unvollständige Paardarbietung ermöglicht. Die Grundlagen des Verfahrens sind bei GULLIKSEN (1956) wie auch bei SIXTL (1967) beschrieben. Bei letztem findet sich auch ein ALGOL-Programm **).

---

*) Das von uns adaptierte Rechenverfahren, Autor M. KOLBE, sah vor, für p = 1.o einen z-Wert von 3.o und für p = o.o einen z-Wert von -3.o zu setzen, was uns zu willkürlich erschien.

**) Leider ist dieses Programm fehlerhaft, wie auch eine Reihe der numerischen Beispiele bei SIXTL (1967). Es empfiehlt sich immer der Rückbezug auf GULLIKSEN (1956).

Es soll nur kurz das Prinzip des Iterationsverfahrens beschrieben werden. Es erstellt im Gegensatz zu den bisherigen Verfahren nicht mit einer einfachen Mittelwertsbildung, ggf. erweitert durch die Varianz als Korrekturgröße, eine Skala. Vielmehr korrigiert es ausgehend von einer ersten Skala (welche im Prinzip als Skalenwerte die Zahlen von o bis n-1 enthalten könnte, zweckmäßigerweise aber von der Skalenlösung nach Fall III ausgeht) diese Skala wiederholend ("Iteration") soweit, bis die Summe der quadrierten Abweichungen der Zellenwerte von empirischer z-Matrix und jeweils reproduzierter z-Matrix so klein wie möglich geworden ist (ein Minimum ergibt). Dabei werden Verteilungen von z = 1.oo zu z = o.oo ausgelassen, gelten also als "nicht beobachtet". In Tab. 4 finden wir für die Untersuchungsgruppe Med 1 alle drei Skalen, nach Fall V, Fall III und Iteration nebeneinandergestellt.

Der Vergleich zeigt, daß die Skalen zum einen verschieden stark gespreizt sind: Bei Fall V werden im Vergleich zur Skala nach der Iteration vor allem die Distanzen am Anfang und am Ende der Skala unterschätzt. Bei Fall III hingegen werden am Anfang und in der Mitte die Distanzen sehr groß, ebenso am Ende. Die Skala ist extrem gespreizt. Des weiteren zeigt sich aber auch eine Umordnung der Rangfolge, vor allem im mittleren Bereich der Skala, ungeachtet der hohen Produkt-Moment-Korrelation (r; Tab. 4, und der etwas besser differenzierenden SPEARMAN-Rang-Korrelation R_s; Tab. 4). Es ist jedenfalls zu folgern, daß das Iterationsverfahren die adäquateste Bestimmungsmethode für die Skalenwerte aus der vorliegenden Distanzmatrix ist.

Dennoch ergibt eine Prüfung der Anpassungsgüte nach dem MOSTELLER-Chi² einen hochsignifikanten Wert $Chi^2 = 236.775$.

Dem Problem der Extremverteilungen haben wir mit dem Iterationsverfahren Rechnung getragen. Als Erklärung

für die mangelnde Güte der Anpassung bleibt immer noch die Möglichkeit, daß die Items nicht eindimensional skalierbar sind. Dagegen spricht allerdings aus den uns bereits bekannten Daten, daß die Anzahl zirkulärer Triaden (intransitiver Urteile) in der Gruppe Med 1 im Durchschnitt erstaunlich gering ist. Bei der Betrachtung des Chi^2-Verfahrens hingegen zeigen sich zwei Probleme, die von BOCK (1958) und vor allem von GULLIKSEN & TUKEY (1958) herausgestellt worden sind. BOCK (1958) kritisiert am MOSTELLER-Verfahren, daß die Güte der Anpassung bei geringerer Anzahl von Objekten eher überschätzt wird. GULLIKSEN & TUKEY (1958) zitieren COCHRAN, der herausstellt, daß das Chi^2-Verfahren bei kleiner Stichprobe von Objekten Unterschiede eher unterschätzt (identisch mit dem Einwand von BOCK), bei sehr großer Stichprobe aber sehr scharf prüft und bereits minimale Differenzen als signifikant erfaßt. Der letzte Kritikpunkt dürfte bei unseren Daten zutreffen. Der Chi^2-Vergleich bezieht sich auf $N\ (N-1)/2 = 15 \times 14/2 = 105$ Tabellenwerte. Er dürfte also zu scharf beurteilen.

Ein weiteres Problem ergibt sich bei der Anwendung auf Fall III und auf die Iteration. MOSTELLER hat das Chi^2-Verfahren nur für Fall V, also unter Annahme gleicher Varianzen entwickelt (und setzt für die aus den winkeltransformierten Matrizen bestimmte Verteilung der Differenzen auch eine nur von N abhängige Größe als Varianz ein). Eine Anwendung auf Fall III und die Iteration dürfte schon deswegen nicht adäquat sein.

GULLIKSEN & TUKEY (1958) schlagen, ähnlich wie BOCK (1958), ein anderes Prüfverfahren vor. Sie prüfen nicht mehr die Güte der Anpassung, was für das Iterationsverfahren ja auch redundant ist, sondern sie betrachten analog zu einer Varianzanalyse drei Komponenten der winkeltransformierten Matrizen: Varianz, die

im linearen Modell aufgeht (S_L; Summe der quadrierten Abweichungen der theoretischen Werte vom empirischen Mittel), Varianz, die durch Abweichung vom linearen Modell entsteht (S_D; Summe der quadrierten Abweichungen der theoretischen - gleich linearen - Zellenwerte von den empirischen Zellenwerten) und der gesamten Varianz (S_T; Summe der quadrierten Abweichungen der empirischen Matrix vom Mittel; SIXTL 1967).

Aus diesen Varianzkomponenten lassen sich zwei Quotienten bestimmen, hier in Anlehnung an GULLIKSEN & TUKEY (1958) mit Rho und Rss bezeichnet, welche beide als Index der (linearen) Skalierbarkeit dienen und analog zu einem Determinations-Koeffizienten zu interpretieren sind. Sie geben an, wie gut die Skala die Gesamtvarianz der empirischen Daten erfaßt (SIXTL 1967). Die Koeffizienten unterscheiden sich dadurch, daß bei Rho direkt die lineare Varianzschätzung wie auch die Varianzschätzung, die sich aus der geschätzten Abweichung vom linearen Modell ergibt, miteinbezogen wird (grob: lineare Varianz mit Abweichungsvarianz zu Gesamtvarianz), während bei Rss allein die Differenz zwischen Gesamtvarianz und Abweichungsvarianz ins Verhältnis zur Gesamtvarianz gesetzt wird. Die Werte sind üblicherweise deswegen nicht identisch, weil die Quadratsummen sich nicht genau additiv zusammensetzen (aufgrund einer non-linearen Verzerrung der winkeltransformierten Verteilung der z-Werte besonders bei großen z-Werten). Als Kriterium für hinreichende Erfüllung der linearen Skalierbarkeit betrachten wir eine numerisch deutlich gegen 1.o (vollkommene lineare Skalierbarkeit) gehende Größe und eine nicht zu große Differenz zwischen beiden Indices.

Rho spiegelt die Häufigkeit von Extremverteilungen 1.o zu o.o. Wenn solche Verteilungen häufiger werden, geht der "Coefficient of Agreement" (KENDALL), den wir aber weniger in den Mittelpunkt stellen, gegen 1.o.

D. h. die Beurteiler stimmen völlig überein, haben alle die gleiche Rangskala für die Krankheiten, aber die Intervallskalenqualität geht verloren, weil sich die Reaktionsverteilungen nicht mehr überlappen, und Rho wird in Relation zu Rss kleiner. Anders gesagt, die Objekte liegen auf der psychologischen Dimension so deutlich auseinander, daß niemand sie mehr "verwechselt", was allerdings weiterhin bedeutet, daß damit das "THURSTONE-Modell" hinfällig ist.

Die Bestimmung der Indices Rho und Rss für die Skala nach dem Iterationsverfahren ergibt für die Gruppe Med 1 Werte von .816 bzw. .961. Damit erweist sich die Beurteilung von Krankheiten im Paarvergleich nach ihrer subjektiv empfundenen Schwere ("wie schlimm, wenn Sie persönlich betroffen wären") als recht gut bis sehr gut linear skalierbar und öffnet den Weg zu einer psychologischen Interpretation.

3.2 Die Auswahl der Krankheitsbilder und die Konstruktion der Paarvergleichsliste

3.2.1 Exkurs: Zum Problem der Anwendung des Krankheitsbegriffes auf psychische Phänomene

Wenn in dieser Untersuchung die psychischen Krankheiten nicht nur einfach neben somatische Krankheitsbilder gestellt werden, sondern die Zuordnung zueinander sogar zur Frage gestellt ist, so wird damit ein inzwischen verbreitetes Vorverständnis zuerst einmal übernommen. Es hat sich inzwischen weitgehend auch im Alltagsverständnis durchgesetzt, bestimmte psychische Ausnahmezustände als Krankheiten anzusehen (wenn es doch nicht so ist, müßte sich das in den Ergebnissen dieser Untersuchung zeigen).

Es ist aber festzustellen, daß dies das Ergebnis einer historischen Entwicklung ist. Und nach wie vor ist der

Begriff "psychische Krankheit" und die Zuordnung zu somatischen Krankheiten problematisch, mehr als für das Alltagsverständnis noch für eine definitorische Analyse aus klinisch-psychologischer und psychiatrischer Sicht. Unsicher sind übrigens auch die Abgrenzungen, was bei bestimmten Erscheinungen Krankheit ist und was nicht.

Im Rahmen dieser Untersuchung ist es unmöglich, auch nur den Versuch einer Lösung für diese Frage zu unternehmen. Die verschiedenen Aspekte sollen aber wenigstens angesprochen werden.

Im Rahmen einer psychologischen Untersuchung liegt es nahe, zuerst die Diskussion anzusprechen, die vor allem im Bereich der Psychologie geführt wurde (obwohl sie von einem Psychiater ausgelöst wurde). Sie schließt sich an an den Artikel von SZASZ (1960) und wurde im deutschen Sprachraum von KEUPP (1972, 1979) rezipiert. SZASZ (1960) versucht zu begründen, daß die Bezeichnung psychischer Abweichungen als "Krankheiten" im Grunde der Ausbau einer ursprünglich metaphorischen Ausdrucksweise zu einem "Mythos" mit nachweisbarer gesellschaftlicher Ausgrenzungs- und Beherrschungsfunktion gegenüber den Betroffenen sei. Die Etikettierung sogenannten "abweichenden Verhaltens" als Krankheiten wird zwar als Fortschritt gegenüber einer Deutung als Zeichen dämonischer Besessenheit angesehen. Es werde damit aber die Ausgrenzung des sich anders Verhaltenden aufrecht erhalten. Sie habe die Aufgabe, solche Störungen nicht auf die gesellschaftliche Ordnung durchschlagen zu lassen. Dabei wird gerade die Ordnung der Gesellschaft aber angeschuldigt, in Wahrheit alle psychischen Störungen bzw. "abweichendes Verhalten" zu produzieren.

KEUPP (1972; 1979) faßt die Kontroverse in die alternative Formulierung "medizinisches Modell" contra "sozialwissenschaftliches Modell". Diese Deklarierung

impliziert zugleich die Definition der Psychologie als "Sozialwissenschaft" sowie deren Zuständigkeit für diese Probleme, für "abweichendes Verhalten", und vice versa die Unzuständigkeit der Medizin, also der Psychiatrie (als somatisch - biologisch - naturwissenschaftlich gesehen). Entsprechende Implikationen - in Kürze sind sie gar nicht vollständig zu analysieren - ergeben sich für die Beurteilung der adäquaten Behandlung psychischer Störungen, wobei in letzter Konsequenz festgestellt wird, daß eigentlich nicht diejenigen, welche sich abweichend verhalten, sondern die Gesellschaft behandelt werden müsse.

Gekennzeichnet ist diese Auffassung weiterhin zumeist dadurch, daß zwischen verschiedenen Formen psychischer Störungen kaum oder gar nicht differenziert wird, selbstverständlich auch nicht zwischen verschiedenen möglichen Ätiologien, da die einzig mögliche ja per definitionem festliegt. Gewisse Ausnahmen werden gemacht bei Störungen, bei denen die körperliche Begründung absolut nicht zu verleugnen ist, etwa bei Folgen eines offensichtlichen Hirntraumas. Insgesamt läßt sich erschließen, daß als Prototyp "abweichenden Verhaltens" eine produktive paranoide Schizophrenie vorschwebt. Die medikamentöse Behandlung ist entsprechend den Prämissen abzulehnen; es wird zwischen verschiedenen Psychopharmaka auch gemeinhin nicht differenziert, alle werden mehr oder weniger als Tranquilizer gedeutet. Aber auch psychotherapeutische Interventionen fallen zumeist unter das Verdikt, nur die Anpassung an eine inhumane Gesellschaftsform erreichen zu wollen, und nicht eine Änderung dieser Gesellschaft (s.o). In popularisierten Formen dieser Richtung gehen die Vorstellungen dahin, daß die "psychisch Kranken" sich dadurch befreien, daß gerade sie als Speerspitze einer Bewegung zur Revolutionierung der Gesellschaft antreten.

Unsere Darstellung ist überspitzt. Festzuhalten bleibt

aber, daß das Echo im deutschsprachigen Bereich zumindest innerhalb der Klinischen Psychologie in psychiatrischen Kliniken wie im Bereich der so heftig attackierten Klinischen Psychiatrie wenig nachhaltig war. Das erklärt sich wohl daraus, daß die überzogene Einseitigkeit dieser Position in der tatsächlichen Arbeit mit dem betroffenen Personenkreis psychisch Kranker in ihrer irrealen Übersteigerung eines bereits vorher von vielen anderen gesehenen richtigen Kerns sehr schnell deutlich wird.

Von heute noch sehr nachhaltiger Wirkung war hingegen die Bemühung von K. SCHNEIDER (1939) um eine definitorische Klärung des Begriffs "psychische Krankheit". Entscheidend ist dabei die Festlegung "Krankheit ist immer körperlich". Dieser Krankheitsbegriff wird als "streng medizinischer" (K. SCHNEIDER 195o; 1967) gesehen. "Psychische" Krankheiten sind demnach alle tatsächlich durch eine somatische Störung verursachten Symptome bzw. Syndrome, also zum einen die exogenen Psychosen, bei denen diese Verursachung erklärt ist, wie dann auch die endogenen Psychosen, für die eine somatische Verursachung zu postulieren ist (heute mit noch mehr gesicherten Befunden als damals). Psychopathologisch sind nach K. SCHNEIDER diese Krankheiten gekennzeichnet durch eine qualitative Abnormität psychischer Erscheinungen, z.B. die "Symptome ersten Ranges" bei schizophrenen Erkrankungen. Ihr gemeinsames Kennzeichen ist, daß die "Sinngesetzlichkeit des Lebens" zerrissen ist (K. SCHNEIDER 1951; 1959 S. 21). Hiervon zu unterscheiden sind die "abnormen Spielarten menschlichen Wesens", abnorme Persönlichkeiten und abnorme (Erlebnis-)Reaktionen - nach K. SCHNEIDER sind auch die abnormen Reaktionen immer wesentlich mitbedingt durch eine abnorme Persönlichkeitsstruktur -, welche nur quantitativ, gradmäßig, entweder in ihrer Intensität oder in ihrer Dauer oder beides, abnorm sind. Letztere können nicht als Krankheiten angesehen

werden, sind auch nicht medizinisch zu behandeln - höchstens unterstützend -, sondern durch Psychotherapie, welche also ausdrücklich nicht zur Medizin gezählt wird (K. SCHNEIDER 195o; 1967). Den Begriff "Neurose" lehnt K. SCHNEIDER explizit ab, schon wegen seiner vom Wortsinn her irreführenden Bedeutung. Offen bleibt die Einordnung psychogener somatischer Störungen.

Mit dem Problem, was eigentlich "seelisch abnorm" ist und warum, wie auch mit dem Problem des Psychogenen setzt sich sehr ausführlich und bis heute nicht in der Ernsthaftigkeit seines Bemühens aufgenommen MÜLLER-SUUR (195o) auseinander. Festzustellen ist, daß er eindeutig in der psychiatrischen Tradition von K. SCHNEIDER steht. Zu seiner zentralen Fragestellung erarbeitet MÜLLER-SUUR in einer phänomenologischen Analyse eine differenzierte Betrachtung verschiedener Normbegriffe im Bereich der Psychologie. Wertende und wertfreie Normbegriffe, also Idealnormen und Realnormen, werden gegeneinander gestellt, dann ersetzt bzw. überschritten durch die Explikation der Begriffe einer "Seinsnorm" und einer "Werdensnorm", jeweils in einer individuellen und in einer "kollektiven" Norm (d. h. die Beziehung des einzelnen zu den anderen Menschen betreffend, nicht etwa nur im heutigen Sinne sozial) ausgeformt. Damit will MÜLLER-SUUR (195o) - auf dem Hintergrund der philosophischen Anthropologie BRENTANOs und deren zentraler Lehre von der Intentionalität des Menschen - ein dynamisches Element im Sinne einer beständigen Entwicklung des Individuums in den Blick bringen. Das Wesen des Seelischen ist danach als Spannung zu verstehen zwischen dem, was jemand derzeit ist und was ihm möglich ist zu werden, zugleich aber - das ist das mit Kollektiv gemeinte - in seiner Beziehung und auch Beeinflussung durch die mitmenschliche Umwelt, welche neben die Beziehung zu sich selbst tritt. Dann versucht MÜLLER-SUUR (195o) eingehend zu begründen, daß abnorm und krankhaft bzw.

Abnormität und Krankheit keineswegs identische Dinge sind. Er kommt - weniger direkt als K. SCHNEIDER, auf viel verschlungenerem Weg - zum gleichen Ergebnis: Psychische Krankheiten sind in einem organischen oder psychischen Prozeß begründet und sind zu erkennen an qualitativ abnormen psychischen Erscheinungen. Dabei steht das Postulat des "psychischen Prozesses" für die endogenen Psychosen, welche in ihrem psychischen Erscheinungsbild den organischen näher stehen. Die Abnormität beruht darauf, daß infolge einer Krankheit die "individuelle Werdensnorm" nicht (mehr) erfüllbar ist. Das Abnorme ist insbesondere darin zu erkennen, daß die Erlebnisweise des Kranken qualitativ verändert ist, was dem Zerreißen des Sinnzusammenhanges entspricht; so hat K. SCHNEIDER es gefaßt. Insgesamt handelt es sich um einen somatogenen psychischen Mechanismus.

MÜLLER-SUUR (195o) fügt dann aber noch einen Bestimmungsversuch des sogenannten "psychogenen psychischen Mechanismus" an. Auch hier kann nämlich psychische Abnormität resultieren, und zwar durch eine in dem Sinne gestörte Intentionalität, daß die individuelle und die kollektive Werdensnorm nicht mehr aufeinander abgestimmt werden können. Mit unseren Worten gesagt: Es resultiert also so etwas wie eine Unangepaßtheit an die Realität. MÜLLER-SUUR spricht von einem Widerspruch zwischen Wollen und Können. Dabei wird ein vereinzeltes Auftreten einer solchen Realitätsunangepaßtheit als durchaus normal angesehen, sozusagen im statistischen Sinne, weil es bei vielen Menschen gelegentlich auftritt. Doch durch eine quantitative Abnormität, also extreme Häufigkeit bei einem einzelnen Menschen, wird sie zur psychogenen Störung. Nur mag MÜLLER-SUUR (195o) dieser psychogenen psychischen Störung - da er sich definitorisch auf den Krankheitsbegriff von K. SCHNEIDER festgelegt hat - nicht den Charakter des Krankhaften zuschreiben. Er führt vielmehr eine Unterscheidung

zwischen Krankheit als dem körperlichen Aspekt und Kranksein als dem seelischen Aspekt ein. So kann er von "psychogenem Kranksein" sprechen, also einem Kranksein mit seinen seelischen Auswirkungen, das nicht in einer Krankheit begründet ist; entsprechend muß es für ihn (es wird nicht mehr deutlich expliziert) auch ein Kranksein, also einen seelischen Reflex, bei körperlichen Erkrankungen geben, und folgerichtig auch ein Kranksein als seelische Auswirkung bei (körperlichen) psychischen Krankheiten.

Mit der Krankheit ist der Mediziner als Vertreter einer Wissenschaft, die heilen will, befaßt; mit dem Kranksein der Arzt als Mensch, welcher helfen will. Die Psychotherapie besteht darin, dem Betroffenen beim psychogenen Kranksein eine Hilfe zur Lösung seines individuellen Normkonflikts zu geben.

Es ist zu ergänzen, daß die Gedankengänge von MÜLLER-SUUR (195o) nur sehr schwer in kurzer Form darzustellen sind. Demgegenüber ist die Position von K. SCHNEIDER wesentlich knapper und klarer. Die Unterschiede liegen z. T. in einem leicht differierenden Ergebnis, vor allem aber darin, daß MÜLLER-SUUR eine eingehende und differenzierte Begründung ausgehend von einer phänomenologischen und psychologischen Analyse versucht.

Es fällt schwer, diese Positionen zu diskutieren, denn auch bei K. SCHNEIDER stehen trotz der apodiktischen Formulierung eingehende Überlegungen auch über psychologische Grundbedingungen im Hintergrund. Dennoch erscheint dem Klinischen Psychologen die Einengung des Krankheitsbegriffes auf körperlich fundierte Störungen unbehaglich. Und bei MÜLLER-SUUR ist vor allem der philosophisch-anthropologische Hintergrund inzwischen fremd geworden, und auch manche Differenzierungen erscheinen artifiziell, insbesondere die abschließende Unterscheidung zwischen Kranksein und Krankheit. Bedeutend vor allem als Anknüpfungspunkt für weiterfüh-

zwischen psychologischer Lebenshilfe und Behandlung gestellt.

Läßt sich hier ein Lösungsansatz skizzieren? Bei allem Ungenügen soll ein Gedanke angesprochen werden. Gerade die Eigenart psychischer Beeinträchtigungen könnte nahe legen, eine definitorische Abgrenzung aus den Ursachen allein zu überdenken. Wesentlich erscheint ein Bedenken dauerhafter Folgen. Man denke etwa an das Beispiel einer schweren Phobie, bei welcher die Beeinträchtigung unmittelbar im psychischen Bereich liegt, etwa in der unbezwingbaren und vom Betroffenen selber als unsinnig angesehenen Angstreaktion vor Hunden. Mittelbare Folgen ergeben sich dann im sozialen Bereich, etwa darin, daß ein Versicherungsvertreter mit einer solchen Hundephobie seinem Beruf nicht mehr adäquat nachgehen kann, weil er ständig befürchten muß, einem Hund zu begegnen. Ein solcher Patient ist in seinem aktuellen beruflichen Leistungsvermögen, z.B. beim Gespräch mit potentiellen Kunden, wenn er immer unter Angst steht, genauso beeinträchtigt, wie es ein anderer mit dem gleichen Beruf sein kann, welcher an einer ausgeprägteren Störung seiner sprachlichen und gedanklichen Flüssigkeit als Residuum nach einer sonst relativ gut ausgeheilten schizophrenen Erkrankung leidet.

Dies Beispiel soll aufzeigen, daß man doch auch wohl von den Konsequenzen her überdenken muß, was eine Beeinträchtigung im Sinne bzw. als Folge einer Krankheit genannt werden muß. Dazu kann man auf den Überlegungen zur Normproblematik von MÜLLER-SUUR (1950) aufbauen, wird aber sicher einiges neu oder anders formulieren. Zugleich wird aber auch deutlich, daß ein Zurückbleiben hinter an sich gegebenen individuellen Möglichkeiten (was mit "individueller Seinsnorm" bzw. "Werdensnorm" umschrieben wurde) und deren Beurteilung wesentlich auch eine Frage an die Psychologie ist.

Damit soll auch begründet werden, warum in dieser Un-

tersuchung körperliche und psychische Krankheiten (gleich welcher Ätiologie) in einen gemeinsamen Kontext zur Beurteilung eingeordnet werden. An den Ergebnissen der Beurteilung durch verschiedene Beurteilergruppen sollte nämlich abzulesen sein, inwieweit psychische Erkrankungen immer noch eine Sonderstellung zugewiesen wird oder in welchem Maß sie in völliger Analogie zu körperlichen Krankheiten etwa unter dem Gesichtspunkt der Beeinträchtigung von Lebensfunktionen gesehen werden.

3.2.2 Auswahl der Krankheitsbilder

Das Hauptinteresse dieser Untersuchung richtet sich auf die psychiatrischen Krankheitsbilder. Uns interessiert, wie vor allem potentiell oder faktisch betroffene Beurteiler den Grad der persönlich empfundenen "Schlimmheit" einstufen.

Die Krankheitsbilder sollten aber nicht für sich isoliert eingestuft werden. Es ist davon auszugehen, daß es eine generelle, wenn auch nur ungefähre Übereinstimmung darüber gibt, daß etwa Grippe eine leichte, Krebs eine hingegen sehr schwere Erkrankung ist. Eine Grippe-Erkrankung dürfte man also entsprechend gelassen hinnehmen, eine Krebs-Diagnose dürfte dagegen als extrem bedrohlich und wahrscheinlich die meisten in große Verzweiflung stürzend erlebt werden. Das ist zu vermuten aufgrund bisheriger Ergebnisse, wie auch der öffentlichen Diskussionen über Krebs und Krebsvorsorge.

Wie steht es hingegen mit anderen Erkrankungen, zumal mit schweren psychiatrischen Krankheitsbildern wie Schizophrenie, aber auch mit organischen Erkrankungen, welche zumindest früher als sozial sehr einschneidend erlebt wurden, wie Tuberkulose als "Arme-Leute-Krankheit", oder einer Geschlechtskrankheit, wie Syphilis. Wie steht dazu wiederum etwa eine vermutlich moralisch

abgewertete Krankheit wie Alkoholismus, oder eine psychische Erkrankung, von der wir aus klinischer Erfahrung am ehesten den Eindruck haben, daß sie bei aller Schwere am meisten sozial akzeptiert ist, nämlich (endogene) Depression?

Aus diesen Überlegungen heraus legten wir zuerst die psychiatrischen Krankheitsbilder fest. Dabei entschieden wir uns, relativ globale Bezeichnungen zu verwenden, von denen zu erwarten war, daß auch Laien sie hinreichend mit Inhalt füllen. Wieweit dies zutrifft, kann nur die Empirie zeigen. Die Bezeichnungen werden auch üblicherweise in den Medien gebraucht. Zumindest im sozusagen "passiven Wortschatz" von Patienten und Angehörigen können sie als vorhanden vermutet werden. Wir wählten als Repräsentanten der klassischen psychiatrischen Gruppe reaktive neurotische Störungen, Sucht, cerebral-organische Störungen, endogene Psychosen (auf Repräsentation von Persönlichkeitsstörungen und Schwachsinnsformen verzichteten wir) folgende Krankheitsbezeichnungen aus:

- Neurose
- Alkoholismus
- Hirnarteriosklerose
  ("Verkalkung" wurde als Erläuterung für die Nichtmediziner hinzugefügt)
- Endogene Depression
  (Erläuterung zu "Endogen" cf. Instruktion)
- Schizophrenie.

Diese fünf Krankheitsbilder sollten eingeordnet werden zusammen mit 1o körperlichen Krankheiten, von leichten bis zu schweren, welche gewissermaßen die zu erstellende Skala erst aufspannen sollten. Da es eine "objektive" Skala der Schwere von Krankheiten nicht gibt, ist eine Auswahl nach verschiedenen, dann auch unterschiedlich ordnenden Kriterien möglich, etwa nach dem Statistischen Jahrbuch (1979) eine Auflistung von Krankheiten nach ihrer Häufigkeit als Todesursache. Diese

Statistik erschien uns weniger geeignet, da sie ein extremes Kriterium für den Schweregrad vor allem in Relation zu den psychiatrischen Erkrankungen enthält, die als Todesursache sicher weniger häufig infrage kommen. Wir orientierten uns an der Statistik der Deutschen Gesetzlichen Rentenversicherung, Band 5o (1978): Rentenzugang und Rentenwegfall im Jahre 1977, und hier wiederum an der Aufstellung "Versichertenrenten wegen Berufsunfähigkeit nach Ursachen, Rentenversicherung der Angestellten - Männer" (eine Gesamtstatistik Angestellte und Arbeiter, Männer und Frauen, ist nicht vorhanden).

Daraus ergab sich folgende ausgewählte Skala, wobei die vorangestellten Zahlenangaben bedeuten: von 1.ooo wegen Berufsunfähigkeit zugegangenen Renten entfielen auf das ausgewählte Krankheitsbild (psychiatrisch relevante Bilder sind eingefügt und unterstrichen):

o.3 Syphilis und sonstige Spirochätenkrankheiten
o.3 Appendicitis
2.3 Querschnittslähmung
3.5 Schizophrenie
5.2 Tuberkulose
16.5 Diabetes mellitus
16.8 Sonstige Psychosen
25.8 Neurosen, Psychopathien und sonstige nichtpsychotische seelische Störungen
52.9 Bösartige Neubildungen (Einzelgruppen summiert)
83.2 Bluthochdruck
92.6 Hirngefäßkrankheiten
151.o Ischämische Herzkrankheiten

Man erkennt, daß auch diese Aufstellung für unseren Zweck nicht direkt brauchbar ist. Was nicht so wichtig ist, aber angemerkt werden soll: Suchterkrankungen tauchen nicht als eigene Kategorie auf (auch in den von uns übergangenen restlichen Gruppen). Sie sind sicherlich für eine Berentung relevant und wohl unter der

Gruppe "sonstige nicht-psychotische Störungen" subsumiert.

Wir haben für unsere Untersuchung folgende Bezeichnungen für die Krankheiten festgelegt:

Syphilis
Blinddarmentzündung
Querschnittslähmung
Schizophrenie
Zuckerkrankheit
Endogene Depression
Neurose
Alkoholismus
Krebs
Bluthochdruck
Hirnarteriosklerose ("Verkalkung")
Herzinfarkt

und eine Krankheit ohne vermutete Relevanz für Berentung ergänzt:

Grippe

als Annäherung eines Nullpunktes. Ferner wurde, da keine Hautkrankheit, die vermutlich psychologisch einen anderen Aspekt bietet, repräsentiert war,

Akne

als häufige Erkrankung des jüngeren Lebensalters hinzugenommen.

3.2.3 Erstellung des Untersuchungsinstrumentes zum Paarvergleich und Modus der Vorgabe

Die klassische Item-Präsentation beim Paarvergleich (GUILFORD 1954), vor allem im Bereich der Psychophysik, geschieht in der Weise, daß sukzessiv jeweils ein Paar von Reizen nach dem anderen, für jeden neuen Probanden in neuer Zufallsreihenfolge, dargeboten wird, was bei physikalischen Stimuli, etwa Gewichten, notwendig ist. Entsprechend wird dann auch bei verbalen Stimuli verfahren, indem etwa jedes Paar auf eine Karte geschrieben wird, die dann sukzessiv dargeboten

werden. Das Verfahren hat den Vorteil, durch die immer neue Zufallsanordnung für jeden Beurteiler evtl. Reihungseffekte auszuschließen. Zugleich ist es aber für größere Untersuchungsserien viel zu zeitaufwendig.

Aus Gründen der Praktikabilität haben wir eine Art Fragebogen oder besser: Paarvergleichsliste erstellt, indem wir eine quadratische Matrix der Krankheiten erstellt haben, wobei Spalten und Zeilen in folgender gleicher Reihenfolge durch die Krankheitsbilder bezeichnet sind (in Klammern die von uns in den Tabellen und Abbildungen gebrauchten Abkürzungen):

Grippe	(Grip)
Akne	(Akne)
Neurose	(Neur)
Blinddarmentzündung	(Blin)
Bluthochdruck	(Blut)
Alkoholismus	(Alko)
Zuckerkrankheit	(Zuck)
Syphilis	(Syph)
Hirnarteriosklerose ("Verkalkung")	(Hirn)
Tuberkulose	(Tb)
Herzinfarkt	(Herz)
Endogene Depression	(Depr)
Querschnittslähmung	(Quer)
Krebs	(Krbs)
Schizophrenie	(Schi)

Die Zellen der oberen Halbmatrix (15 x 14/2 = 1o5) symbolisieren dann die möglichen Paare.

Die Anordnung der Paarvergleichsliste erfolgte in zwei Schritten:

1. wurden, im Allgemeinen diagonal durchgehend in Richtung der die Matrix nicht symmetrisch teilenden Diagonalen, die Paare in der Weise aufgelistet, daß

die Liste zu sieben Seiten mit 15 Paaren geordnet wurde, wobei jedes Krankheitsbild auf einer Seite generell nur zweimal erscheinen sollte (auf einer von sieben Seiten jeweils dreimal), und

2. wurde die Stellung der Paarlinge zueinander durch Los bestimmt, um eine Reaktionstendenz (etwa häufiger den ersten Paarling anzukreuzen) möglichst auszugleichen.

Der Paarvergleichsliste wurde als Titelblatt eine schriftliche Instruktion vorangestellt (vgl. Anhang 2), die im wesentlichen die Entscheidungsregel erklärte, darauf hinwies, daß Gleichheitsurteile nicht gestattet seien und dann noch eine Erläuterung zur Bezeichnung der Krankheitsbilder gab, die für die Medizinstudenten und für die anderen Gruppen etwas differierte.

Die Beurteiler wurden gleichermaßen instruiert, von ihrer eigenen Vorstellung bzw. Kenntnis einer Krankheit auszugehen. Rückfragen im Sinne "Was genau meinen Sie damit?", und vor allem bei den Medizinstudenten der an sich berechtigte Einwand "Die Krankheiten können sich nach ihren Unterformen oder auch Stadien sehr unterscheiden" wurde insofern versucht vorwegzubegegnen, indem eingeräumt wurde "Manche Bezeichnungen könnten Ihnen etwas weitgefaßt vorkommen". In einer zusätzlichen mündlich gegebenen Instruktion (die gesamte Instruktion wurde vorgelesen und bei Rückfragen durch Wiederholung bzw. Vorführen an einer Tafel bzw. Papier erläutert), wurde den Medizinstudenten erläutert, daß die Bezeichnungen der Krankheiten deswegen so gewählt wurden, weil beabsichtigt sei, die gleiche Untersuchung auch mit Nicht-Medizinern durchzuführen, daß deswegen also eher Syndrombilder als genaue diagnostische Kategorien gewählt wurden und die Bezeichnungen zumeist nach dem allgemein üblichen Sprachgebrauch erfolgten (Hirnarteriosklerose wurde nur für die Nicht-Mediziner durch den Zusatz ("Verkalkung") ergänzt).

Die anderen Untersuchungsgruppen hatten mit der weiten Terminologie keine Probleme, hier wurde vielmehr bei drei Begriffen üblicherweise nach der genaueren Bedeutung gefragt: Am häufigsten nach dem Wort "endogen", dann nach "Akne", dann (erst) nach "Neurose". Allen Gruppen wurde hierzu mündlich folgende einheitliche Erläuterung gegeben: ""Endogen" ist ein medizinisches Fachwort, das deswegen hier steht, weil wir die Vergleiche auch Medizinern vorgelegt haben, die das ja immer etwas genauer wissen wollen. Es bedeutet "von innen heraus kommend". Sie wissen aber doch, was "Depression" ist. Denken Sie daran, was man üblicherweise als Depression kennt. "Endogen" ist für Sie nicht so wichtig."

Hier wurde von uns natürlich die Voraussetzung gemacht, daß üblicherweise eine allgemeine Vorstellung von Depression da ist, die im wesentlichen nicht differenziert zwischen reaktiver bzw. neurotischer Depression und psychotischer Depression, sondern eher die Symptomatik sieht. Wir sind uns bewußt, daß diese Voraussetzung nicht unproblematisch ist und bei einer Interpretation der Ergebnisse beachtet werden muß. Andererseits erschien uns eine zweifache Ausführung von Depression, getrennt als "reaktiv (neurotisch)" und "endogen (psychotisch)" als ungünstig, da wir davon ausgehen, daß diese Unterscheidung zu viel Verwirrung stiften würde.

Zu Neurose: "Neurose ist eine seelische Krankheit." Bei weiterem Nachfragen: "Ich möchte Ihnen das jetzt nicht noch genauer erklären, dann würde ich vielleicht Ihr persönliches Urteil beeinflussen. Es kommt mir ja darauf an, Sie mit den Vorstellungen urteilen zu lassen, wie unsere Patienten auch üblicherweise am Anfang zu uns in die Behandlung kommen."

Zu Akne: "Akne ist eine Hautkrankheit vor allem bei jungen Leuten."

Ein weiterer wesentlicher Unterschied der Vorgabe für Medizinstudenten einerseits und die verschiedenen übrigen Untersuchungsgruppen andererseits bestand darin, daß letztere explizit schriftlich und mündlich, und zwar wiederholt, aufgefordert wurden, sich vorzustellen, wie "Sie persönlich urteilen, wenn Sie davon betroffen wären." Für die Mediziner hingegen wurde das bewußt offen gelassen; sie wurden im nachhinein gebeten anzumerken, ob sie eher objektiv als Arzt beurteilt hätten oder sich eher versucht hätten, in die Situation eines Kranken hineinzuversetzen. Dann sollten sie entsprechend ein "A" (Arzt) oder "B" (Betroffener) auf dem letzten Blatt notieren. Dabei bestand jedoch die Gruppe bei der Durchführung der Untersuchung darauf, eine dritte Möglichkeit eingeräumt zu erhalten, ein "S" - "Ist für mich schwierig" - notieren zu dürfen. Vor der Bearbeitung des Paarvergleiches wurden dann noch Angaben zu persönlichen Daten erbeten: Bei den Medizinstudenten Anfangsbuchstabe des Vornamens, Geburtsdatum (um feststellen zu können, welche Listen bei der geplanten Wiederholung am Ende des Semesters vom gleichen Beurteiler stammen), Geschlecht, Alter, Semesterzahl. Bei den übrigen Gruppen: Alter, Geschlecht, Angaben zur Schul- und Berufsbildung (als grobe Indikatoren des Bildungsniveaus).

3.3 Bedeutung des Informationsstandes über die Krankheiten und der persönlichen Kriterien für die Beurteilung

3.3.1 Informiertheit über die Krankheitsbilder

3.3.1.1 Kursus in Psychiatrie für die Medizinstudenten

Bei einer Einstufung von Krankheiten nach dem Grad ihrer "Schlimmheit" ist vermutlich die kognitive Komponente von entscheidender Bedeutung. Bisherige Untersuchungen, die unserem Konzept vergleichbar sind, zeigen das sehr deutlich.

Bei den Medizinstudenten bot sich die Möglichkeit, "Informationsstand" als quasi unabhängige, wenn auch nicht voll kontrollierbare und natürlich nicht manipulierbare Variable insofern einzuführen, daß die Skalierung der Krankheiten zweimal, zu Beginn und zu Ende des Kurses in Psychiatrie erbeten wurde. So lassen sich bei den psychiatrischen Krankheiten anhand potentieller Veränderungen der Skala Wirkungen vermehrter Kenntnis abschätzen.

Der Unterricht in Psychiatrie an der Medizinischen Fakultät der Universität des Saarlandes besteht im (verpflichtenden) Kern aus dem Praktikum der Psychiatrie in kleinen Gruppen mit Patientenvorstellung sowie einer mehr systematischen Ergänzungsvorlesung, ebenfalls konkretisiert durch Vorstellung von Patienten. Dabei werden insbesondere im Praktikum zu ungefähr gleichen zeitlichen Teilen vor allem fünf Syndromgruppen berücksichtigt: reaktive/neurotische Störungen (und Persönlichkeitsstörungen), Suchterkrankungen, cerebral-organische Bilder, (endogene) affektive Psychosen und Schizophrenien. Man kann in summa davon ausgehen, daß der Informationsstand der Studenten zum Ende des Kurses hin gestiegen und detaillierter ist, und daß sie die Krankheitsgruppen besser differenzieren können. Der Pflichtteil kann fakultativ ergänzt werden durch zwei Zusatzvorlesungen, deren eine systematische Fragestellungen der Psychiatrie abdeckt, vor allem eine Systematik der allgemeinen Psychopathologie gibt, dann u. a. inhaltliche Aspekte der klinischen Psychologie und Psychotherapie in der Psychiatrie, der forensischen Psychiatrie und der Psychopharmakotherapie erörtert. Die zweite Zusatzvorlesung stellt vorzugsweise ausgewählte Krankheitsbilder unter besonderer Berücksichtigung des Krankengutes einer Landesnervenklinik vor.

Dabei kommt es jedoch vor, daß eindrucksmäßig wie auch von der zeitlichen Akzentuierung her manchmal bestimmte psychiatrische Bilder besonders in den Vordergrund

treten. Soweit wir das für das betreffende Semester kontrollieren konnten, lernten die Studenten besonders eindrucksvolle Krankheitsbilder mit endogenen (monopolaren) Depressionen kennen, während sich die Demonstration schizophrener Erkrankungen am Schluß des Semesters etwas zusammendrängte (und bei der zweiten Befragung nicht ganz abgeschlossen war).

Bei den körperlichen Krankheiten setzen wir einen hohen Informationsstand der Studenten - zumindest in Relation zu den übrigen Gruppen medizinischer Laien - im zweiten Teil des klinischen Abschnittes ihres Studiums voraus. Einflüsse, welche Veränderungen zwischen den beiden Untersuchungszeitpunkten in Bezug auf somatische Krankheiten bewirken könnten, können wir hier nicht beurteilen.

3.3.1.2 "Wissenstest" über Krankheiten

Auch bei den übrigen Gruppen war eine Erfassung der Kenntnis über Krankheiten erforderlich. Eine ideale Lösung wäre gewesen, einen Wissenstest über die ausgewählten 15 Krankheitsbilder zu konstruieren und an einer repräsentativen Stichprobe zu eichen. Empirisch gesicherte Ergebnisse darüber zu erhalten, wie zutreffend Kenntnisse über Krankheiten in der Bevölkerung sind, wäre ein sehr wichtiges Unterfangen von grundsätzlicher Bedeutung. Es ist aber bisher nicht geleistet worden. Nur wäre im Rahmen dieser Arbeit der Aufwand für die Konstruktion eines solchen Tests, der dann 15 eigenständige Untertests für jede einzelne Krankheit enthalten müßte, übergroß und nicht zu leisten. Und für die Beurteiler hätte die zusätzliche Bearbeitung von hinreichend langen und so erst reliablen 15 Subtests eine zusätzliche zeitliche und damit motivationale Belastung bedeutet.

Wir begnügten uns daher mit der Konstruktion eines informellen "Tests", der für uns im wesentlichen zwei Kriterien erfüllen sollte:

1. Er sollte für eine von zwei Kontrollgruppen ohne manifeste Krankheitsbelastung vor allem die Funktionen des Lerninstrumentes haben. Dazu war notwendig, einen minimalen Schwierigkeitsgrad zu erreichen.
2. Für alle Gruppen sollte er eine sehr grobe Kontrollmöglichkeit des Wissens über die zu beurteilenden Krankheiten bieten, um eklatante Inkompetenz durch völlig mangelnde Kenntnis prüfen zu können. Dazu war wiederum ein nur geringer Schwierigkeitsgrad hinreichend.

Es wurde nach der Vorlage eines bereits recht differenzierten, aber populärwissenschaftlichen Handbuches über Krankheiten ("Wie funktioniert das? - Der Mensch und seine Krankheiten", Bibliographisches Institut Mannheim 1973) ein Lerntest entworfen*), in dem zu jedem Krankheitsbild drei zutreffende Feststellungen formuliert wurden. Und zwar in der Art, daß durch möglichst eindeutige Kombination von Einzelinformationen das Krankheitsbild in wichtigen Aspekten beschrieben wurde (vgl. Anhang 3). Die drei richtigen Items wurden zu einem kleinen Untertest für jede der 15 Krankheiten komplettiert, indem zwei weitere für eines der anderen Krankheitsbilder richtige Feststellungen, welche für die hier abgefragte Krankheit aber recht leicht als falsch zu erkennen waren, in jeweils zufällig wechselnder Abfolge hinzugefügt wurden.

In der Instruktion (vgl. Anhang 3) wurde einmal darauf hingewiesen, daß man nur einen groben Überblick über den Kenntnisstand bekommen wolle. Mündlich wurde außerdem ergänzt, daß niemand sich "blamieren" könne, weil man ja nicht für den einzelnen herausfinden wolle und könne, wie genau er informiert sei. Zur weiteren Vereinfachung wurde folgende Lösungsregel angegeben und

---

*) Wir haben einigen psychiatrischen Kollegen für ihre hilfreiche Beratung zu danken.

schriftlich wie mündlich eingehend erläutert: Immer drei Feststellungen seien zutreffend,richtig, die beiden anderen falsch. Es wurde in Kauf genommen, daß damit die Items eines Krankheitsbildes ihre stochastische Unabhängigkeit verloren, also auch eine grobe Reliabilitätsprüfung vereitelt wird (die aufgrund der geringen Itemzahl pro Krankheit aber an sich fragwürdig gewesen wäre; LIENERT 1969). Geprüft wurde einzig die Schwierigkeit der Items (Anhang 3; nach der einfachsten Formel für Zuordnungsaufgaben). Fernerhin wurde ein Gesamtrohwert für jeden Beurteiler bestimmt (vgl. unten Tab. 16). Ebenso wurde auf eine Trennschärfenanalyse verzichtet, da hier ohne Belang.

3.3.2 Erfassung der Kriterien der Beurteilung

SÜLLWOLD (1967) hat darauf hingewiesen, daß von wesentlicher Bedeutung nicht nur die Skalierung von Objekten durch eine Gruppe ist, sondern auch die Beweggründe, welche sie zu ihrer spezifischen Einstufung veranlassen. Eine sinnvolle Interpretation einer erstellten Skala hängt also auch von der Kenntnis dieser Kriterien ab.

Für die Medizinstudenten haben wir diese zweite Einflußgröße dadurch zu erfassen gesucht, daß wir sie baten, auf einem angefügten Blatt (vgl. Anhang) neben der Kennzeichnung ihres Beurteilungsaspektes (A, B oder S; s.o.) möglichst der Rangfolge nach die Kriterien ihrer Entscheidung zu notieren. Diese Kriterien konnten dann nach ihrer Häufigkeit, bei vorhergehender Zusammenfassung unter inhaltlichem Aspekt in größere Gruppen, ausgezählt werden.

Für die übrigen Untersuchungsgruppen, bei denen wir einen größeren Zeitverbrauch für die Bearbeitung des Paarvergleichs erwarteten sowie eine weitere zeitliche Belastung durch den informellen Wissenstest, verzichteten wir auf eine genauere schriftliche Befragung. Denn untersuchungsmethodisch stand im Vordergrund, die Moti-

vation der freiwilligen Teilnehmer und damit die Qualität der Bearbeitung des Paarvergleichs nicht zu beeinträchtigen. Wir boten den einzelnen Gruppen vielmehr eine Möglichkeit zur nachfolgenden Erklärung und Diskussion der gesamten Untersuchung sowie zur Beantwortung von Informationsfragen zu den Krankheiten oder zur Psychologie allgemein an. Jeweils ein größerer Teil der Gruppen nahm dieses Angebot gerne wahr. Dieses Gespräch bot dann die Gelegenheit, möglichst zu Beginn (der manchmal sehr eingehenden, vor allem bei den beiden Patientengruppen auch sehr von der Erörterung ganz persönlicher Erfahrungen geprägten Gespräche) die verbleibende interessierte Gruppe nach ihren Beurteilungskriterien zu fragen. Sie wurden anschließend in einem Gedächtnisprotokoll festgehalten und standen in dieser Form für die Stützung der Interpretation zur Verfügung.

3.4 Kennzeichnung der Beurteilergruppen und Schema der Untersuchung

3.4.1 Charakteristik der Untersuchungsgruppen

Die endgültige Zusammenstellung der Untersuchungsgruppen erfolgte nach einer Prüfung der Kompetenz der Beurteiler. Dabei war das erste Kriterium die Vollständigkeit der Bearbeitung der Paarvergleichsliste. Hier mußten in fast allen Gruppen Beurteiler ausgeschieden werden (Tab. 5), bei der Gruppe der Tuberkulosepatienten (Tb) sogar eine erhebliche Anzahl.

Das zweite Kriterium bestand in der Prüfung der individuellen Urteilskonsistenz nach dem KENDALL-Verfahren für die Anzahl zirkulärer Triaden und dem von uns schärfer formulierten Kriterium, nur bis 7o solcher intransitiver Urteile zuzulassen. Daraufhin schied nur noch ein Beurteiler der Gruppe Kontrolle 2 aus, der dieses Kriterium weit überschritt (Tab. 5).

Die erste der Untersuchungsgruppen haben wir bereits er-

Tabelle 5:

Prüfung der Kompetenz der Beurteiler nach korrekter Bearbeitung der Paarvergleichsliste und hinreichender Urteilskonsistenz (= Anzahl der intransitiven Urteile; "zirkuläre Triaden")

	Untersuchungsgruppen					
	Med 1	Med 2	Tb	Alk	Kontr 1	Kontr 2
Paarvergleich vorgelegt bei N =	67	37	7o	48	69	76
ausgeschieden	1	o	22	7	8	4
verbleibendes N =	66	37	48	41	61	72
Zirkuläre Triaden:						
o	4(6%)	3(8%)	2(4%)	1(2%)	1(2%)	2(3%)
1 - 1o	45(68%)	19(51%)	29(6o%)	3o(75%)	31(51%)	36(5o%)
11 - 2o	15(23%)	13(35%)	11(23%)	9(22%)	18(3o%)	19(26%)
21 - 7o	2(3%)	2(5%)	6(13%)	1(2%)	11(18%)	14(19%)
71 u. mehr	o	o	o	o	o	1(1%)
ausgeschieden	o	o	o	o	o	1
endgültig verbleibendes N=	66	37	48	41	61	71

wähnt. Es handelt sich um 66 Medizinstudenten (Tab. 5) vorzugsweise des 8. bis 1o. Semesters, also in der klinischen Ausbildung, welche im Sommersemester 198o am Kursus in Psychiatrie der Medizinischen Fakultät der Universität des Saarlandes in Homburg/Saar teilgenommen hatten. Die Gruppe wurde in erster Linie gewählt aufgrund ihrer Verfügbarkeit für eine erste Erprobung der Untersuchungsmethode im Blick auf Praktikabilität, Zeitverbrauch und vor allem eindimensionale Skalierbarkeit der Krankheiten.

Da die Gruppe gut brauchbare Ergebnisse erbrachte, wurde sie in die Hauptuntersuchung mit aufgenommen. Die Liste wurde in der 3. Vorlesungsstunde des Semesters zum erstenmal vorgelegt.

Die erste Vorgabe der Paarvergleichsliste wurde dann ergänzt durch eine zweite Vorgabe in der drittletzten Vorlesungsstunde des Semesters. Diese zweite Beurteilergruppe, Med 2, ist zum größten Teil (29 Beurteiler) identisch mit der Gruppe Med 1, hinzu kamen aber 8 Studenten, die an der Erstuntersuchung nicht teilgenommen hatten, so daß sich für Med 2 ein N = 37 ergibt. Beide Gruppen bestanden zu etwa einem Viertel bis einem Drittel aus Studentinnen, zu Dreiviertel bis zwei Drittel aus Studenten. Bei einem medianen Alter von 25.o6 bzw. 23.45 Jahren reichte die Altersverteilung grob gesehen über das Altersjahrzehnt der Zwanzigjährigen. Die meisten Beurteiler studierten jeweils im 8. Semester (Tab. 6).

Weiterhin nahmen wir zwei Gruppen von Patienten, also Betroffene in die Untersuchung auf: Repräsentanten für eine organische Erkrankung und für eine psychiatrische Erkrankung. Das Auswahlprinzip war einmal pragmatisch: Erreichbarkeit in einer größeren homogenen Gruppe, um nicht auf Einzelvorgaben angewiesen zu sein. Das führte zur Wahl einmal einer Gruppe mit Lungentuberkulose, zum anderen einer Gruppe von Alkoholkranken (Tab. 7),

Tabelle 6:
Charakteristik der Untersuchungsgruppen Medizinstudenten 1 (Med 1) und Medizinstudenten 2 (Med 2) nach Geschlecht, Alter, Studiensemester (absolute (f) und relative (f%) Häufigkeiten)

	Med 1 (N = 66)		Med 2 (N = 37)	
	f	f%	f	f%
weiblich	25	38	9	24
männlich	41	62	28	76
Alter				
2o	o	o	1	3
21	1	2	1	3
22	9	14	6	16
23	13	2o	1o	27
24	5	8	4	11
25	9	14	6	16
26	9	14	3	8
27	7	11	1	3
28	5	8	3	8
29	4	6	o	o
3o u. älter	4	6	2	5
Median:	25.o6		23.45	
im 8.	43	65	26	7o
9.	1o	15	2	5
1o.	11	17	8	22
höheren	2	4	1	3
Semester				
Modalwert:	8. Semester		8. Semester	

Tabelle 7:
Charakteristik der Untersuchungsgruppen Tuberkulose-Patienten (Tb) und Alkoholkranke (Alk) nach Geschlecht, Alter, Schul- und Berufsausbildung (absolute (f) und relative (f%) Häufigkeiten)

	Tb (N = 48)		Alk (N = 41)	
	f	f%	f	f%
weiblich	33	69	17	59
männlich	15	31	24	41
Alter				
18-2o	4	8	o	o
21-3o	11	23	12	29
31-4o	8	16	13	32
41-5o	7	15	12	29
51-6o	11	23	4	1o
61 u. älter	7	15	o	o
Median:	42.oo		37.oo	
Schulausbildung (erreichter Abschluß)				
Hauptschule	3o	63	2o	49
Real-,höhere Sch.	7	15	9	22
Fachschule	8	17	6	15
Abitur	2	4	4	1o
ohne Schulabschl.	1	2	o	o
Sonstiges	o	o	o	o
keine Angabe	o	o	o	o
Modalwert:	Hauptsch.		Hauptsch.	
Berufsausbildung				
angelernt	1o	21	4	1o
Lehrberuf	29	6o	27	66
Studium	2	4	3	7
ohne	6	13	6	15
noch in Ausbildung	1	2	o	o
keine Angabe	o	o	1	2
Modalwert:	Lehrberuf		Lehrberuf	

für die jeweils entsprechende Fachkliniken bzw. Fachabteilungen zur längerfristigen Behandlung eingerichtet sind. Ein zweites Auswahlkriterium war, daß Lungentuberkulose bzw. Alkoholismus das alleinige vorherrschende Krankheitsbild sein mußte, also ohne weitere ernsthafte Begleit- oder Folgeerkrankungen. Beide Gruppen sollten die konkrete Erfahrung mit einem Krankheitsbild repräsentieren; gedacht war dabei vor allem daran, ob sich eine solche Erfahrung nach dem Modell eines "Ankerreizes" auswirkt und die Einstufung der Krankheiten insgesamt oder die Einstufung der eigenen Erkrankung verändert.

Beide Gruppen, die Tuberkulosepatienten (Tb) und die Alkoholpatienten (Alk), sind homogen in Bezug auf das jeweilige Krankheitsbild. Hingegen ist die Altersverteilung der Gruppe Tb sehr heterogen (Tab. 7). Die Gruppe setzt sich ferner aus Untergruppen aus insgesamt vier Kliniken bzw. Fachabteilungen zusammen. Weiterhin zeigte sich, daß im Vergleich zu anderen Gruppen (Tab. 7 und Tab. 8) eine erhebliche Anzahl von Teilnehmern an der Untersuchung ausscheiden mußte, weil sie die Paarvergleichsliste nicht korrekt oder nicht vollständig bearbeitete, also offensichtlich trotz mehrfacher Erklärung die Instruktion nicht verstand. Bei einigen Teilnehmern war das schon im Verlauf der Bearbeitung festzustellen. Da sie aber sehr gerne mitmachen wollten, wurden sie selbstverständlich nicht zurückgewiesen. Die Zahl der nicht-verwertbaren Bearbeitungen in der Gruppe Alk ist dagegen geringer (Tab. 7).

Schließlich wurde die Untersuchung ergänzt durch zwei Gruppen von Rekruten der Bundeswehr, Kontrollgruppe 1 (Kontr 1), N = 61, und Kontrollgruppe 2 (Kontr 2), N = 71, die als Repräsentanten einer normalgesunden Bevölkerungsgruppe gewählt wurden (Tab. 8). Beide Gruppen bestanden ausschließlich aus männlichen Probanden, waren außerdem noch mehr altershomogen als die Gruppen Med 1 und Med 2; andererseits zeigten sie eine breitere

Tabelle 8:
Charakteristik der Untersuchungsgruppen Kontrollgruppe 1 und 2 nach Geschlecht, Alter, Schul- und Berufsausbildung (absolute (f) und relative (f%) Häufigkeiten; Prüfung der Unterschiede auf Signifikanz mit Chi²)

	Kontr 1 (N = 61)		Kontr 2 (N = 71)	
	f	f%	f	f%
weiblich	-	-	-	-
männlich	61	1oo	71	1oo
Alter				
18	1	2	o	o
19	2o	33	12	17
2o	23	38	33	46
21	11	18	1o	14
22	4	7	6	8
23 u. älter	2	4	5	7
keine Angabe	o	o	5	7
Median:	19.91		2o.o6	
Chi² bei df = 4 : 5.79 (n.s.)				
Schulausbildung (erreichter Abschluß)				
Hauptschule	25	41	28	39
Real-,höhere Sch.	12	2o	7	1o
Fachschule	4	7	5	7
Abitur	17	28	24	34
ohne Schulabschl.	1	2	2	3
Sonstiges	1	2	o	o
keine Angabe	1	2	5	7
Modalwert:	Hauptsch.		Hauptsch.	
Chi² bei df = 4 : 3.66 (n.s.)				
Berufsausbildung				
angelernt	4	7	1	1
Lehrberuf	36	59	38	54
Studium	o	o	4	6
ohne	1o	16	16	23
noch in Ausbildung	2	3	1	1
keine Angabe	9	15	11	15
Modalwert:	Lehrberuf		Lehrberuf	
Chi² bei df = 5 : 7.o5 (n.s.)				

Streuung in Bezug auf Schul- und Berufsbildungsniveau.

Die beiden Kontrollgruppen unterschieden sich voneinander nicht in Hinblick auf Alter, Bildungsniveau und Geschlecht. Der Unterschied zwischen beiden Gruppen liegt einzig in der Anordnung der Vorgabe von Paarvergleichsliste und "Wissenstest". Während Kontr 1 wie die beiden Patientengruppen Tb und Alk zuerst die Paarvergleichsliste zu bearbeiten hatten, erhielt Kontr 2 zuerst den Wissenstest über die Krankheiten als Lerninstrument und dann unmittelbar folgend die Paarvergleichsliste.

Bei einer eingehenden Betrachtung der Charakteristika der sechs verschiedenen Beurteilergruppen (Tab. 6, 7 und 8) stellt sich die Frage, wieweit die Ergebnisse, welche aus ihren Beurteilungen resultieren, zu verallgemeinern sind. Eine Generalisierbarkeit von Aussagen hängt prinzipiell von der Repräsentativität solcher Gruppen für eine Grundgesamtheit ab.

Es ist festzuhalten, daß den Gruppen keine streng empirische Repräsentativität zukommt. Wohl aber kann man von einer "logischen" Repräsentativität ausgehen, allerdings nicht bei allen Gruppen in gleichem Maß. Alle Gruppen sind nicht Zufallsstichproben aus wohldefinierten Populationen; für eine Verallgemeinerbarkeit von Ergebnissen ist solch eine strenge Prozedur wünschenswert, aber aus Gründen des Aufwands bei psychologischen Untersuchungen fast nie erreicht. Mit "logischer" Repräsentativität ist gemeint, daß die quasi "natürliche" Zusammensetzung einiger unserer Gruppen der Ziehung einer Zufallsstichprobe zwar nicht gleichkommt, aber ähnelt. Das dürfte für die Medizinstudenten wie für die beiden Kontrollgruppen gelten. So dürften die Medizinstudenten mit gewissen Einschränkungen als repräsentativ für alle deutschen Medizinstudenten in diesem Abschnitt ihres Studiums gelten, wobei zwei einschränkende Bedingungen herausgehoben werden sollen: Einmal ist

möglich, daß an diesem Studienort solche Studenten überrepräsentiert sind, deren Heimatort nicht allzuweit entfernt ist. Gewichtiger dürfte aber sein, daß eine andersartige Ausbildung, vor allem inhaltlicher Prägung, an anderen medizinischen Fakultäten wahrscheinlich ist. Dabei ist anzunehmen, daß sich das vor allem auf eine spezifische Bewertung psychischer Faktoren bei der Beurteilung der Krankheiten auswirken könnte. Es ist denkbar, daß an verschiedenen Orten psychologische, psychosomatische und psychiatrische Gesichtspunkte durchaus verschieden und vor allem mit verschiedenem Gewicht akzentuiert werden.

Noch weniger einschränkende Faktoren sind bei den beiden Gruppen der Bundeswehrrekruten anzunehmen. Besondere Auswahlgesichtspunkte für die Zuordnung zu einem Ort sind nicht zu erkennen. Abgesehen davon, daß einzig selbstverständlich die Gruppe der Zivildienstleistenden nicht vertreten ist (bei denen man möglicherweise an eine akzentuiertere Einstellung zu psychischen Erkrankungen denken könnte), kann man diese Gruppen als repräsentativ für eine altersgleiche männliche Gesamtpopulation ohne offensichtliche gesundheitliche Probleme ansehen.

3.4.2 Schema der kontrollierten und variierten Einflußgrößen

Die verschiedene Auswahl der Gruppen wie auch die Variation der Aufgabenstellungen garantiert eine Vergleichsmöglichkeit der Ergebnisse in Abhängigkeit von als relevant zu vermutenden Einflußgrößen. Es sind selbstverständlich nicht alle relevanten Einflußgrößen erfaßt, wohl aber wesentliche. Das sind einmal der Grad der Kenntnis über die zu beurteilenden Krankheiten, weiter der allgemeine Bildungsstand, der zumindest für Kontr 1 und Kontr 2 als gleich anzusehen ist. Ferner die persönliche Erfahrung einer Krankheit (Gruppen Tb und Alk), und schließlich die in einer Beurteilungs-

gruppe vorhandenen individuellen Kriterien für die Bildung der persönlichen Rangfolge der Krankheiten nach dem Grad ihrer subjektiv empfundenen Schwere.

Weiterhin war nachträglich eine Prüfung auf geschlechtsspezifische Beurteilungseinflüsse (bei hinreichend großen Untergruppen) möglich , indem Med 1 und Alk in weibliche und männliche Beurteiler aufgeteilt und die Skalen getrennt bestimmt wurden. Insgesamt ergibt sich das Schema der Tab. 9, welches einen abschließenden Gesamtüberblick über den Untersuchungsplan vermittelt.

Tabelle 9:
Schematische Darstellung des Untersuchungsplanes

	Gruppen					
kontrollierte bzw. geprüfte Variablen	Kontr 1	Kontr 2	Tb	Alk	Med 1	Med 2
Beurteilungs-konsistenz	+	+	+	+	+	+
Bildungsstand	+	+	+	+	+	+
altersspezif. Einflüsse	-	-	-	-	-	-
geschlechtsspe-zif. Einflüsse	-	-	-	+	+	-
"Wissenstest"	+	(+)	+	+	-	-
variierte Bedingungen						
"Wissenstest" als Lerninstrument	-	+	-	-	-	-
eigene Erfahrung	-	-	+	+	-	-
Medizin-Stud.	-	-	-	-	+	+
spezieller Psychiatrie-unterricht	-	-	-	-	-	+

(Symbole: + = ja; - = nein)

4 Ergebnisse

4.1 Die Skalierung durch die Gruppen der Medizinstudenten (Med 1 und Med 2)

Bei der Darstellung der Logik und des Berechnungsverfahren der Paarvergleichsmethode sind bereits - als Beispiel - Ergebnisse der Gruppe der Medizinstudenten dargestellt worden. Damit wurde vor allem begründet, weshalb es bei einer Beurteilung der Ergebnisse sinnvoll ist, sich auf die durch das Iterationsverfahren erstellte Skala zu beziehen. Weiterhin war zu begründen, daß zu einer Beurteilung der Güte der Anpassung der Skala an die empirischen Daten nicht der MOSTELLER-Chi^2-Test, sondern die Indices der Skalierbarkeit nach GULLIKSEN & TUKEY heranzuziehen sind. Demnach sollen auch im weiteren für alle Beurteilergruppen die Skalen nach Fall V und Fall III wie auch die geschätzten Streuungswerte für Fall III und die Chi^2-Werte des MOSTELLER-Tests für jede Skalierungsmethode, ferner die Interkorrelation der Skalen nach den verschiedenen Skalierungsverfahren (Produkt - Moment - Korrelation r) und KENDALLs "Coefficient of Agreement" dargestellt werden.

Die Skalen der Krankheiten, die sich aus der Beurteilung durch die Gruppe Med 1 ergeben haben, finden sich in Tab. 1o, eine graphische Darstellung bereits für alle Gruppen gemeinsam in Abbildung 1. Zur weiteren Interpretation wird aber ausschließlich die jeweilige Skala nach dem Iterationsverfahren herangezogen (und deswegen allein in den Abbildungen dargestellt). Sie ist jeweils entsprechend gekennzeichnet (Tab. 1o, "Iteration"). Grippe wird als die Krankheit beurteilt, die am wenigsten "schlimm" ist. Diese jeweils "leichteste" Krankheit haben wir als Anfangspunkt der Skala gewählt, entsprechend der Absicht, die wir bei der Auswahl der Krankheiten hatten. Der Skalenwert o.ooo bedeutet also

Tabelle 1o:

Skalenwerte der Gruppe Med 1

N = 66

Coefficient of Agreement U = o.556

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.4o6	o.ooo	o.ooo
Akne	o.1o3	1.o57	o.429	o.355
Neur	1.123	o.799	1.752	1.784
Blin	o.433	o.889	1.oo3	o.7o8
Blut	1.o16	o.575	1.811	1.454
Alko	1.699	o.641	2.452	2.23o
Zuck	1.232	o.628	1.976	1.66o
Syph	1.413	o.759	2.125	1.957
Hirn	1.755	1.177	2.577	2.578
Tb	1.824	o.726	2.8o4	2.129
Herz	1.9o1	o.693	2.5o3	2.855
Dèpr	1.534	o.816	2.257	2.186
Quer	2.398	2.123	4.441	3.473
Krbs	2.883	1.717	5.o51	3.61o
Schi	2.o86	o.995	2.984	2.882
Chi^2:	749.o68		91o.427	236.771
S_D :	11.334		13.794	3.589
S_L :	61.495		59.99o	86.1o9
S_T :	1o4.9o2		1o4.9o2	1o4.9o2
Rho :	.5696		.5516	.8156
Rss :	.8753		.8483	.96o5

Korrelationen (Produkt-Moment; r):

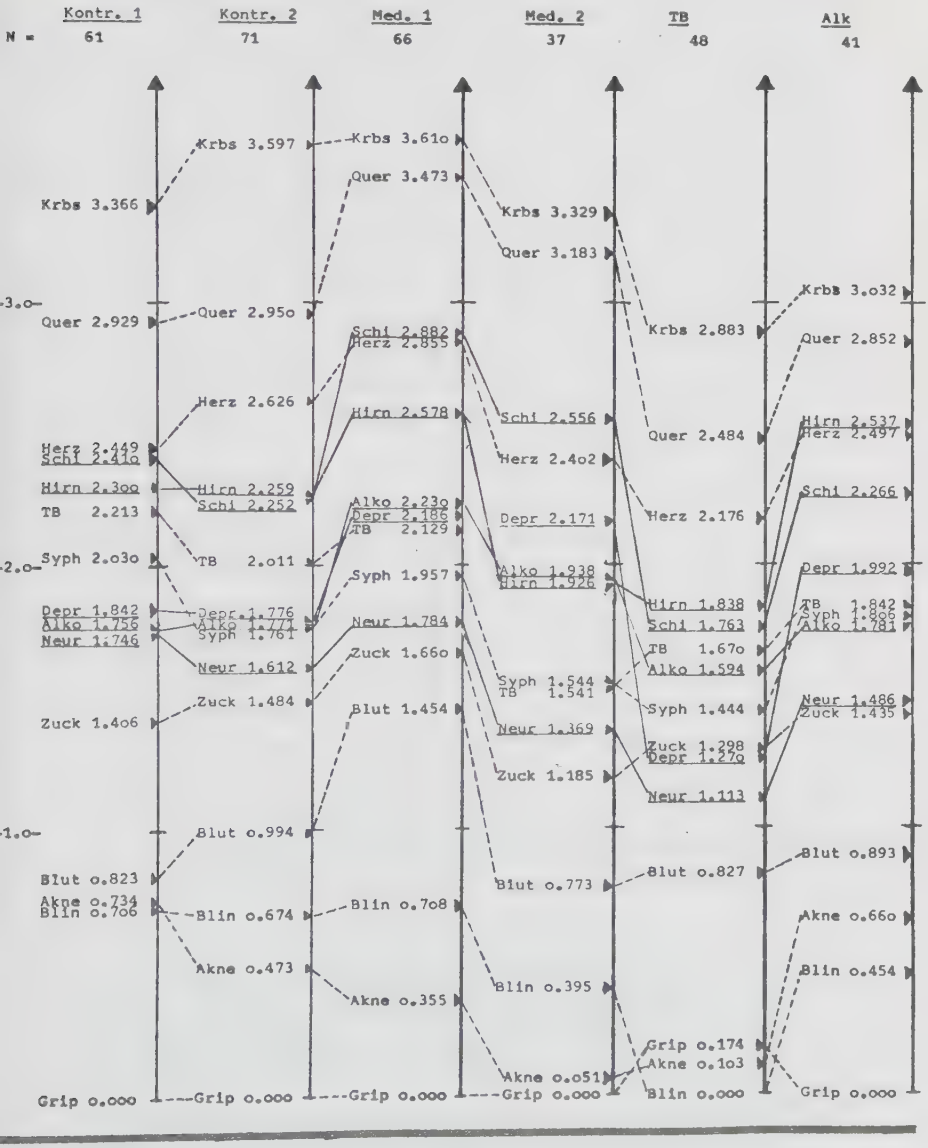
Skala V zu Skala III: .971

Iteration zu Skala V : .986

Iteration zu Skala III: .953

Abbildung 1:

Skalen der sechs Beurteilergruppen zum Vergleich der Skalenwerte (z-Werte) und -positionen der 15 Krankheiten



(Verbindungslinien zwischen den psychiatrischen Krankheitsbildern durchgezogen, sonst gestrichelt zur besseren Kennzeichnung der Veränderungen der Skalenpositionen)

nicht, daß Grippe als gar nicht "schlimm" empfunden wird. In einer anderen Auswahl von Krankheiten oder bei einer anderen Gruppe von Beurteilern könnte Grippe durchaus eine andere Position einnehmen (vgl. später die Gruppe Tb). Als schwerste Erkrankung gilt der Gruppe Med 1 hingegen Krebs mit einem Skalenwert von 3.61o.

Soweit bisher Einstufungen von Krankheiten untersucht worden sind, auch in anderen Auswahlen von Krankheiten und mit anderen Beurteilergruppen und Methoden und Skalierungsverfahren, zeigt sich über alle Untersuchungen hinweg als stabiles Ergebnis: Krebs wird als das Krankheitsbild angesehen, welches am meisten Furcht auslöst, am meisten beschäftigt (salience), für die schwerste Krankheit gehalten wird. Auch für die Gruppe der angehenden Ärzte gilt das.

Zwischen Grippe und Krebs ordnen sich die anderen Krankheitsbilder auf der Skala ein. Im Hinblick auf die oben diskutierte Repräsentativität der hier vertretenen Beurteilergruppen für eine jeweilige Grundgesamtheit stellt sich die Frage nach der Fehlerhaftigkeit der Skalenwerte der einzelnen Krankheiten. Ins Praktische gewendet: Läßt sich für die Skalenwerte ein Meßfehler bestimmen, aufgrund dessen man den wahrscheinlichen Bereich des "wahren" Skalenwertes schätzen kann? Läßt sich dann daraus abgeleitet auch folgern, welche Skalenwerte sich wirklich unterscheiden, welche anderen etwa in ihrer Position auch vertauscht sein könnten?

Zur Frage der Meßgenauigkeit von Skalenwerten aus der Paarvergleichsmethode findet sich in der gesamten Literatur kein Lösungsvorschlag. Theoretisch erscheinen auf dem erreichten Skalenniveau (Intervallskala) zwei Wege denkbar.

Einmal könnte man aus der Schätzung des Standardfeh-

lers des Mittelwertes - die Skalenwerte sind Mittelwerte einer Reaktionsverteilung - entsprechend dem Konzept der "Confidential Limits" Vertrauensgrenzen bestimmen, innerhalb derer sich die "wahren" Skalenwerte bewegen. Üblicherweise wird dabei die z-Verteilung der Standardnormalkurve verwandt. Angemessener wäre, da hier die Gruppen relativ klein sind, die Verwendung der t-Verteilung (FRÖHLICH & BECKER 1971). Beides setzt strenggenommen voraus, daß die Beurteilergruppen Zufallsstichproben aus einer jeweiligen definierten Grundgesamtheit sind. Das Fehlen dieser Voraussetzung ist aber nicht der entscheidende Hinderungsgrund für die Anwendung eines solchen Verfahrens. Vielmehr wird bei diesem Konzept auch ein Streuungswert der Verteilungen benötigt. Diesen Streuungswert hat man zwar bei einer Lösung nach Fall III des "Law of Comparative Judgement", nicht aber für das Iterationsverfahren nach GULLIKSEN. Da aber in allen Fällen die Iteration die bessere Lösung zur Repräsentation der Ausgangsmatrix bietet, entfällt dieses Konzept.

Eine zweite denkbare Möglichkeit wäre, für die jeweilige P-Matrix (relative Häufigkeit des "schlimmer - weniger schlimm" - Urteils für jedes Stimulus-Paar) nach dem Konzept des Standardfehlers von Prozentwerten (CLAUSS & EBNER 1971) zu Vertrauensintervallen für die Parameter pi der "Grundgesamtheit" zu kommen. Man erhielte dann jeweils zwei zusätzliche P-Matrizen für die unteren und oberen Grenzwerte, aus denen entsprechend dem Ursprungsverfahren dann die jeweiligen zugehörigen Skalen zu bestimmen wären. Abgesehen vom Rechenaufwand ist dieses Konzept aber nicht gangbar, weil es viel zu häufig starke Abweichungen von p bzw. q vom Wert o.5 gibt, was zumal bei kleineren Stichproben die Anwendung dieser Methode verbietet.

Da also eine exakte Lösungsmethode nicht vorliegt, die Frage nach den Vertrauensgrenzen aber dennoch von Interesse ist, wenn man auf einer Intervallskala Unter-

schiede von Skalenwerten interpretieren will, wurde versucht, einen dritten ungefähren und grob schätzenden Weg zu beschreiten.

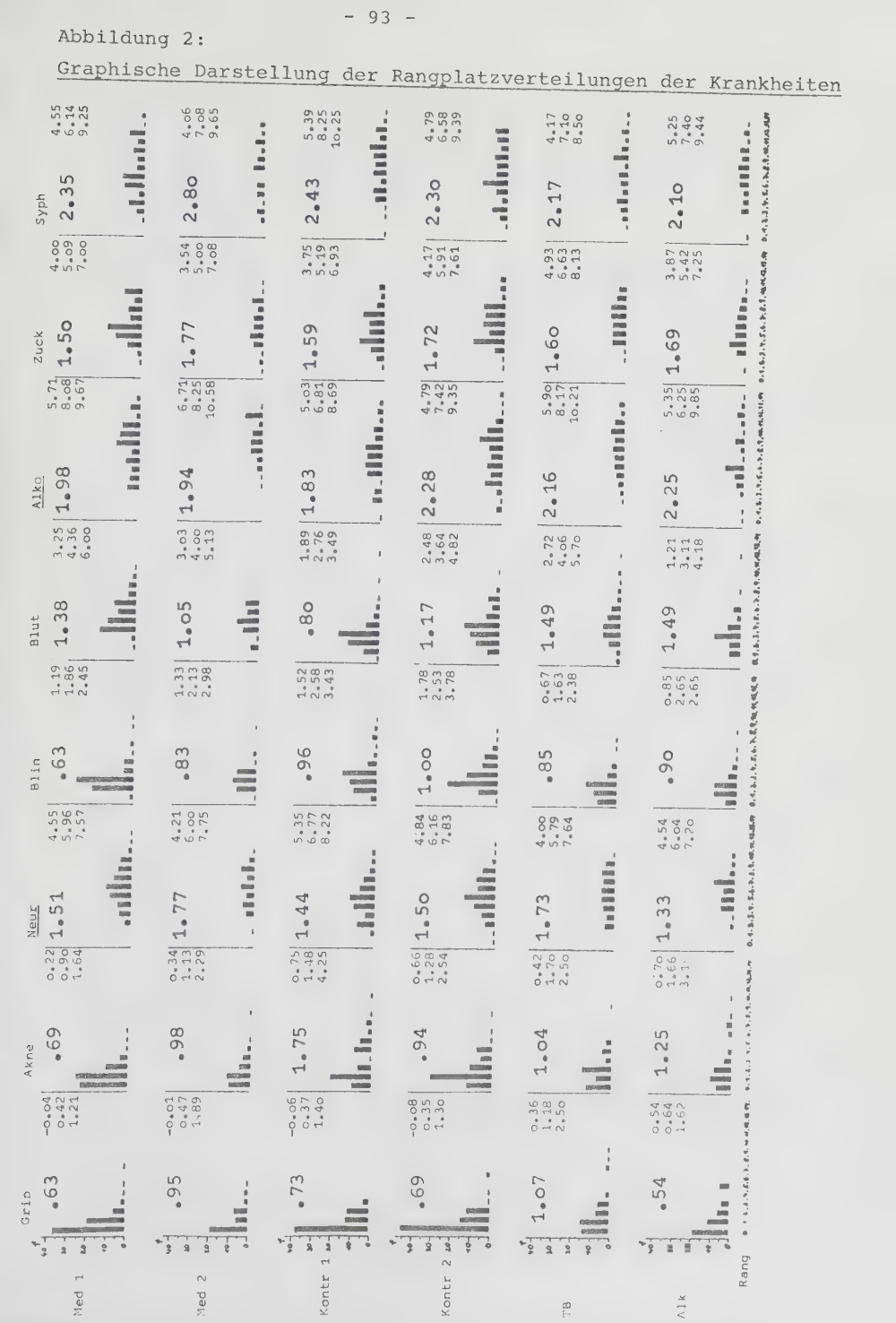
Aus den individuellen Präferenzmatrizen läßt sich für jeden Beurteiler eine Rangreihe der Krankheiten nach subjektiv empfundener Schwere ermitteln. Wenn man für jede Beurteilergruppe für jedes Krankheitsbild die implizit zugeteilten Rangplätze, welche von o bis 14 reichen können (keine bis zu 14 Bevorzugungen), auszählt, ergibt sich eine Häufigkeitsverteilung auf ordinalem Skalenniveau, für welche sich Median und Streuungswerte (Q1 und Q3 bzw. der mittlere Quartilabstand mQ) ermitteln lassen. Diese Rangplatzverteilungen mit den Kennwerten mQ (links in großen Ziffern) und Q1, Median, Q3 (in dieser Reihenfolge von oben nach unten rechts bei jedem Histogramm in kleinen Ziffern) finden sich in Abb. 2. Aus der graphischen Darstellung und den beigegebenen Kennwerten läßt sich zumindest grob erkennen, bei welchen Krankheitsbildern die Beurteiler in den jeweiligen Gruppen breit streuen, bei welchen weniger und wo kaum.

Ein Überblick zeigt, daß die Beurteilung mancher Krankheitsbilder in allen oder fast allen Gruppen, wie die relativ geringe Streuung der zugeteilten Ränge aufweist, recht einheitlich ist. Das gilt z. B. für Grippe, teilweise für Akne, Blinddarmentzündung und Herzinfarkt, ebenso für die Beurteilung von Querschnittslähmung und von Krebs.

Relativ breite Verteilungen finden sich für Neurose, Bluthochdruck, Zuckerkrankheit, Tuberkulose. Uneinheitlich zwischen den Gruppen ist die Urteilsstreuung bei Schizophrenie, teilweise streuen die Rangplatzverteilungen sehr stark, bei Med 1, Med 2 und Kontr 1 aber weniger. Bis auf Neurose, wo einheitlicher geurteilt wird, streuen die Verteilungen für die psychiatrischen Erkrankungen eher weit, wobei Alkoholismus

Abbildung 2:

Graphische Darstellung der Rangplatzverteilungen der Krankheiten



(Kennwerte der Verteilungen: große Ziffern = mittl. Quartilab-

Abbildung 2 (Forts.):

Graphische Darstellung der Rangplatzverteilungen der Krankheiten

immer mit die größten Dispersionen aufweist. Eine vergleichbar durchgängig breite Streuung in allen Gruppen findet sich bei den körperlichen Krankheiten nur für Syphilis.

Aus dieser eher "optischen" Verteilungsanalyse der Urteilsstreuungen soll für die Interpretation von Skalendifferenzen folgende Konsequenz gezogen werden: Geringe Skalendifferenzen, zumal bei Krankheiten, deren Verteilungen generell weit streuen, sollten nicht interpretiert werden. Vielmehr sollte man davon ausgehen, daß die Abfolge bei Krankheiten, deren Skalenwerte geringe Differenzen aufweisen, eher zufällig ist.

Auf eine Gesamtbetrachtung der Skala von Med 1 angewendet heißt das also (Tab. 1o und Abb. 1), daß in der Skalenposition unmittelbar benachbarte Krankheiten nicht unbedingt ihrem subjektiven Schweregrad nach voneinander unterschieden sein müssen. Bei "mittelbaren Nachbarn" hingegen kann man wohl auf einen Unterschied in der Skalenposition vertrauen. Weiterhin gibt es aber Gruppen von Krankheiten, die praktisch identische Positionen haben, also auf der Skala sehr dicht beieinander liegen. Bei Med 1 einmal die Dreiergruppe Tuberkulose, Alkoholismus und endogene Depression, zum anderen die Zweiergruppe Herzinfarkt und Schizophrenie. Die Gruppe Medizinstudenten halten diese Krankheiten also jeweils für etwa gleich schlimm.

Mit gewissen Einschränkungen läßt sich eine Struktur von Krankheitsgruppen beschreiben, deren Grenzen sich bei den drei letzten aber überdecken: eine Gruppe von wenig schlimmen Krankheiten: Grippe, Akne und Blinddarmentzündung. Eine zweite Gruppe von mittelschweren Krankheiten mit Bluthochdruck, Zuckerkrankheit, Neurose, Syphilis, und mit fast gleicher Skalenposition Tuberkulose und endogene Depression und Alkoholismus. Dann eine dritte Gruppe von relativ schweren Erkrankungen: Hirnarteriosklerose und Herzinfarkt sowie

Schizophrenie mit wiederum fast identischen Werten. Schließlich zwei als sehr schlimm eingestufte Krankheiten als vierte Gruppe: Querschnittslähmung und Krebs.

Bei der Beurteilung der Krankheiten besteht eine relativ hohe Übereinstimmung innerhalb der Gruppe Med 1, wie der Coefficient of Agreement U = o.556 anzeigt. Theoretisch könnte er den Wert 1.oo erreichen. Das bedeutete aber, daß alle Zellen der absoluten Häufigkeitsmatrix entweder N (= Beurteiler) oder o enthielten. Dieser Wert ist also als Bezug für eine Beurteilung des Grades der Übereinstimmung nicht realistisch, da bei völliger Übereinstimmung eine Voraussetzung des Skalierungsmodelles, daß es nämlich Überschneidungen der Reaktionsverteilungen gibt, nicht mehr gegeben wäre. Und wie später noch zu sehen sein wird, findet sich bei Med 1 der zweithöchste Übereinstimmungskoeffizient von allen Beurteilergruppen. Die Medizinstudenten haben also insgesamt recht homogene Vorstellungen über die Rangordnung der Krankheiten.

Besonderes Interesse richtet sich auf die Einordnung der psychiatrischen Bilder (in Tab. 1o und Abb. 1). Unter ihnen wird Schizophrenie als die schlimmste Krankheit angesehen, in der Skalenposition gleich mit Herzinfarkt. Dann folgen in ungefähr gleichen Abständen Hirnarteriosklerose, Alkoholismus und endogene Depression in etwa gleicher Position (identisch mit Tuberkulose), und als am wenigsten schlimme psychische Krankheit Neurose. Insgesamt werden die psychiatrischen Krankheitsbilder den mittelschweren bis schweren, nicht unbedingt aber Schizophrenie den sehr schweren Erkrankungen zugeordnet. Neurose liegt ungefähr auf gleicher Distanz zur leichtesten wie schwersten Krankheit, zu Grippe und Krebs.

Daß Schizophrenie für die psychiatrischen Krankheiten die Skala anführt, war zu erwarten, daß Neurose sie

beschließt, ebenfalls. Keine Voraussage erschien möglich, wie sehr Schizophrenie von Krebs abgesetzt wird. Auch daß Querschnittslähmung deutlich näher zur schwersten Erkrankung zugeordnet wird als Schizophrenie, war nicht vorauszusehen. Bemerkenswert ist weiterhin, daß Alkoholismus doch als eher schwere Krankheit angesehen wird und unter den psychiatrischen Erkrankungen eine mittlere Position einnimmt, deutlich von Neurose abgesetzt. Zu bedenken ist, daß den Beurteilern aber vorgestellt wurde, an ein "durchschnittliches" Erscheinungsbild zu denken. Aufgrund der Fragestellung kann ja nicht die Tatsache berücksichtigt werden, daß es insgesamt recht verschiedene Ausprägungsgrade einer Erkrankung, zumal einer psychischen Erkrankung geben kann, und daß etwa eine schwere Neurose den Betroffenen mehr beeinträchtigen kann als einen anderen Patienten eine milde verlaufende oder gut remittierte schizophrene Psychose.

Die Bedeutung einer solchen Skala ist aber für sich alleine genommen recht schwer zu beurteilen. Eine weiterreichende Interpretation erschließt sich erst aus dem Vergleich mit den Skalen anderer Gruppen oder mit Skalen, die unter anderen Bedingungen erstellt worden sind.

Gegenüber der Gruppe Med 1 unterscheidet sich die Gruppe Med 2 darin, daß der erneuten Beurteilung der Krankheiten der oben beschriebene Kursus in Psychiatrie vorausgegangen ist. Die Zahl der Beurteiler hat sich (zu unserem Bedauern) verringert von 66 auf 37. Von den 37 Beurteilern der Gruppe Med 2 haben 29 auch zum ersten Zeitpunkt die Paarvergleichsliste bearbeitet.

Vor einem Vergleich beider Skalen stellte sich die methodische Frage, ob es gerechtfertigt sei, beide Gruppen mit jeweils 66 bzw. 37 Beurteilern einfach gegeneinander zu vergleichen, oder ob sich der Ver-

gleich streng auf die identischen Beurteiler beschränken müsse. Betrachtet man die Frage unter dem ursprünglich beabsichtigten Gesichtspunkt einer strengen Vorher-Nachher-Veränderung, muß man den Vergleich auf die identische Gruppe beziehen. Betrachtet man andererseits beide Gruppen als nicht notwendig parallel, kann man sie in ihrer ursprünglichen Größe nehmen und die Daten von Tab. 1o (Med 1) und Tab. 11 (Med 2) bzw. die Abb. 1 heranziehen.

Da jedoch das aus theoretischen Erwägungen heraus nicht allein entschieden werden kann, haben wir die Skalen für die identischen Beurteilergruppen mit N = 29 gesondert berechnet und sie zum Vergleich zu Abb. 1 in Abb. 3 festgehalten. Es zeigt sich, daß die Skalen für Med 1 und Med/Vorher (N = 29) wie auch für Med 2 und Med/Nachher (N = 29) bis auf geringfügige Abweichungen identisch sind. Die Korrelationen - obwohl als Kenngröße für Übereinstimmung in diesem Fall der Skalen nicht unproblematisch, betragen für Med 1 zu Med 1 (N = 29) r = .995, für Med 2 zu Med 2 (N = 29) r = .997.

Als erster Unterschied der Skalen Med 1 und Med 2 bzw. Vorher-Nachher fällt die geringere Spreizung der Skalen Med 2 (bzw. Med/Nachher) auf. Vom ersten Eindruck her, insbesondere wenn man alle Skalen der Abb. 1 betrachtet, könnte man vermuten, daß diese Tatsache vielleicht nur mit der geringeren Anzahl von Beurteilern zusammenhängt. Wir haben deshalb diese Frage eingehend vor allem an Daten der Gruppe Kontr 2 geprüft. Aus theoretischen Überlegungen folgt, daß die Spreizung der Skala sicher abhängig ist vom Ausmaß der Urteilsdifferenziertheit in einer Gruppe, also davon, wie weit Krankheiten in Bezug auf das Kriterium klar eingestuft werden oder nicht, und damit von der Anzahl der zirkulären Triaden bzw. intransitiven Urteile, welche ein Indikator für diese Unter-

Tabelle 11:

Skalenwerte der Gruppe Med 2

N = 37

Coefficient of Agreement U = o.562

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.674	o.ooo	o.ooo
Akne	o.298	1.556	o.844	o.o51
Neur	1.177	o.57o	2.456	1.369
Blin	o.272	o.967	1.183	o.395
Blut	o.721	o.673	2.o94	o.773
Alko	1.757	o.727	3.22o	1.938
Zuck	1.132	1.o59	2.458	1.185
Syph	1.24o	o.9o2	2.47o	1.544
Hirn	1.5o6	o.667	2.689	1.926
Tb	1.359	o.543	2.821	1.541
Herz	1.883	o.783	3.16o	2.4o2
Depr	1.622	o.78o	2.76o	2.171
Quer	2.216	o.912	3.511	3.183
Krbs	2.762	2.11o	5.932	3.329
Schi	2.oo6	1.o78	3.43o	2.556
Chi^2 :	468.812		6o4.627	173.717
S_D :				4.697
S_L :				85.839
S_T :				1o6.543
Rho :				.7989
Rss :				.9491

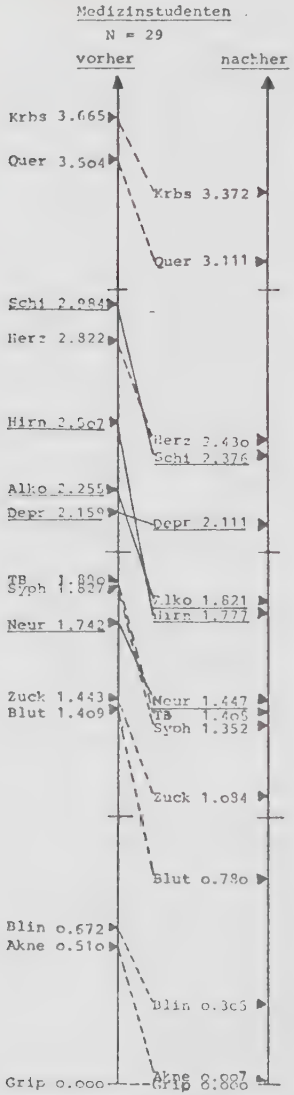
Korrelationen (r):

Skala V zu Skala III: .956

Iteration zu Skala V : .986

Iteration zu Skala III: .915

Abbildung 3:
Skalen (Iteration) der identischen Beurteilergruppen aus Med 1 und Med 2



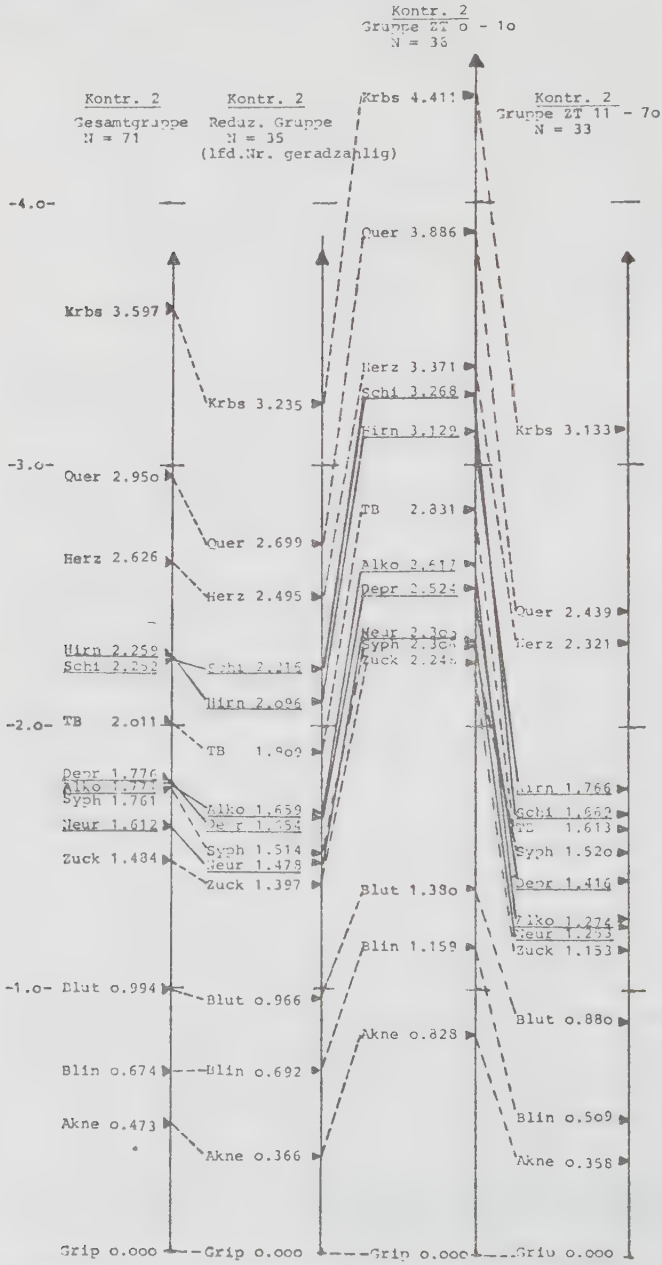
scheidungsfähigkeit sind (wir setzen einen konstanten Katalog von Krankheiten voraus). Zum zweiten ist die Spreizung der Skala abhängig von der Anzahl der extremen Verteilungen (d. h. von Proportionen, die deutlich von o.5 : o.5 abweichen). Je mehr Extremverteilungen (und je höher damit der Coefficient of Agreement U), desto gespreizter wird die Skala. Im Äußersten wird sie so weit gespreizt, daß keine z-Werte mehr zugeordnet werden können (wenn in einer Spalte der Häufigkeitsmatrix nur noch 1oo % bzw. o % vertreten sind) und sich die Skalierbarkeit verliert.

Von der Größe der Beurteilergruppe kann die Spreizung der Skala dagegen nicht abhängen. Es gibt nämlich keinen Zusammenhang zwischen der Gruppengröße und der Auftretenswahrscheinlichkeit von extremen Reaktionsverteilungen einerseits wie der Auftretenswahrscheinlichkeit von intransitiven Urteilen andererseits.

In Abb. 4 sind unter diesem Gesichtspunkt die Ergebnisse einer Aufteilung der Gruppe Kontr 2 unter verschiedenen Gesichtspunkten dargestellt. Es findet sich einmal die Skala der Gesamtgruppe mit N = 71, welche identisch ist mit der Skala Kontr 2 in Abb. 1. Dann findet sich eine Skala für eine nach Zufall reduzierte Gruppe mit N/2 = 35 (Auswahlprinzip war eine geradzahlig laufende Nummer), weiter zwei Skalen für Untergruppen, welche nach der Anzahl intransitiver Urteile gebildet wurden, und zwar einmal die Beurteiler, welche o bis 1o zirkuläre Triaden aufwiesen, und die Restgruppe, die mehr als 1o zirkuläre Triaden aufwies, wobei letztere wesentlich weniger homogen in Bezug auf das Kriterium "Anzahl intransitiver Urteile" ist. Sehr deutlich ist die unterschiedliche Spreizung der Skala in den beiden letzten Gruppen festzustellen.

Wenn man diese Zwischenüberlegung nun auf den Vergleich der Gruppen Med 1 und Med 2 anwendet, ergibt

Abbildung 4:
Skala der Gruppe Kontr 2 und Skalen von Untergruppen aus Kontr 2 (jeweils Iteration)



sich, daß die Gruppe Med 2 weniger differenziert. Eine Erklärung für dieses Phänomen könnte sein, daß durch Zunahme an Information nicht nur eine stärkere Differenzierung verursacht sein könnte, sondern im Gegenteil bei Überschreitung eines bestimmten Maßes an Information wiederum eine stärkere Verunsicherung und damit geringere Differenzierung. Weitere Erklärungsmöglichkeiten müssen offen bleiben.

Für die Frage nach der Einordnung psychiatrischer Erkrankungen interessiert besonders eine Veränderung dieser entsprechenden Skalenwerte in Relation zur Veränderung der Skalenwerte körperlicher Erkrankungen zwischen den Gruppen Med 1 und Med 2. Bei ausschließlicher Betrachtung der körperlichen Erkrankungen ist als erstes festzustellen, daß aufgrund der Stauchung der gesamten Skala bei Med 2 die Skalenwerte generell gegenüber Med 1 reduziert sind. Bei den somatischen Erkrankungen bleibt aber die Reihenfolge (bzw. Rangfolge) insgesamt gleich (auch bei den identischen Gruppen mit N = 29), wenn man von einem Wechsel zwischen Tuberkulose und Syphilis absieht. Dazu ist aber zu ergänzen, daß ja im großen und ganzen die Skalenpositionen zumindest benachbarter Krankheiten aufgrund eines anzunehmenden Meßfehlers auch vertauscht sein könnten bzw. als gleich betrachtet werden können. Es ist also die im wesentlichen gleiche Reihenfolge organischer Krankheiten herauszuheben, während die Reihenfolge der psychiatrischen Krankheiten sich zwischen Med 1 und Med 2 verändert hat. So behält endogene Depression trotz gestauchter Skala seine Position bei, wird also relativ schwerer beurteilt, Hirnarteriosklerose hingegen wird sozusagen über Veränderung durch die Stauchung hinaus relativ leichter gesehen. In Bezug auf die Bewertung von endogener Depression ist hier zu ergänzen, daß aufgrund eher zufälliger Einflüsse die Darstellung dieses Krankheitsbildes in der besagten Unterrichtsveranstaltung durch

sehr eindrucksvolle und schwere konkrete Fälle ein großes Gewicht erhielt. Das ist ein Beleg für die Bedeutsamkeit von Ausmaß und Inhalt von Information über Krankheiten.

Die numerischen Abstände jeweils zwischen Krebs, Querschnittslähmung, Schizophrenie, Alkoholismus bleiben bei Med 1 und Med 2 ungefähr gleich, der numerische Abstand von Alkoholismus zu Neurose vergrößert sich. Man kann also weiterhin folgern, daß auch Neurose in Relation leichter gesehen wird.

Insgesamt ist festzustellen, daß der Unterricht in Psychiatrie einen deutlichen Einfluß auf die Beurteilung der psychischen Krankheiten hat. Es erfolgen von Med 2 gegen Med 1 vor allem Umbewertungen innerhalb dieser Krankheitsbilder. Es resultiert aber nicht, daß psychische Erkrankungen nun generell in Relation zu körperlichen Erkrankungen als schwerer oder leichter eingestuft werden.

4.2 Die Kriterien der Medizinstudenten für die Einstufung der Krankheiten

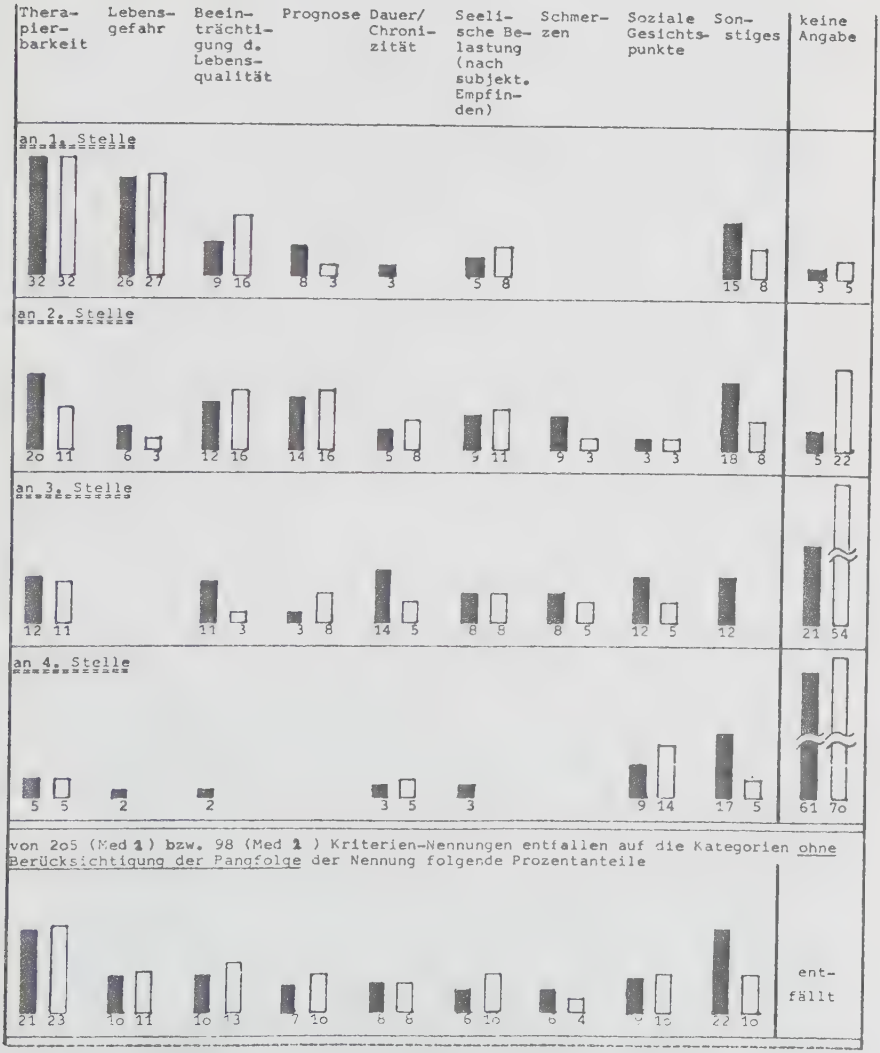
Bei der Gruppierung der von den Medizinstudenten in freier Antwortform zu notierenden Kriterien benutzten wir zur Kennzeichnung der Kriteriengruppen die Formulierungen, welche von den Studenten zumeist mehrfach selber gebraucht wurden. Die Darstellung der Ergebnisse in Form von Histogrammen findet sich in Abb. 5. Die Beurteiler konnten fünf und mehr Kriterien der Rangfolge der Wichtigkeit nach nennen. Die Mehrzahl der Gruppe Med 1 beschränkte sich auf bis zu 3 Kriterien, die Gruppe Med 2 auf bis zu 2 Kriterien.

Zwischen verminderter Nennung von Kriterien bei Med 2 und der geringeren Spreizung der Skala ist ein Zusammenhang anzunehmen. Wahrscheinlich ist, daß bei der Wiederholung der Befragung die Motivation zur Mitar-

Abbildung 5:

Relative Häufigkeiten und Rangfolge der von Med 1 und Med 2 genannten Kriterien für die Beurteilung der Krankheiten

(Zeilen: Rangfolge; Spalten: Kriterien-Kategorien; Zahlen unter den Säulen: % der Gruppe nennt die Kategorie; bei Zeile 5: % aller Nennungen; Med 1: gefüllte, Med 2: offene Säulen)



beit gesunken war, daß für die Medizinstudenten bereits ein gewisser Sättigungseffekt eintrat. Dafür spricht auch die vermehrte Häufigkeit intransitiver Urteile als Indikator für weniger große Sorgfalt. Wie aus Tab. 5 zu ersehen ist, ist der relative Anteil von Beurteilern mit mehr als 1o zirkulären Triaden bei Med 2 größer als bei Med 1 (allerdings bei einer Chi^2-Prüfung nicht signifikant; $Chi^2 = 2.91$ bei $df = 3$).

Wie schlimm eine Krankheit sei, wird von den Medizinstudenten beider Gruppen vor allem nach zwei Gesichtspunkten beurteilt: Einmal danach, wie weit sie als therapierbar, behandelbar, heilbar angesehen wird (Med 1: 32 %, Med 2: 32 %), dann danach, wie weit sie lebensgefährlich ist, tödlich ist, wie hoch die Wahrscheinlichkeit eines "letalen Ausgangs" ist (Med 1: 26 %, Med 2: 27 %). Dann folgt der Häufigkeit der Nennungen nach unter den an erster Stelle genannten Kriterien "Beeinträchtigung der Lebensqualität". Gemeint ist damit eine Einschränkung vieler oder aller "normalen" Lebensvollzüge, also die mittelbaren Folgen einer Erkrankung, welche sich in einer mehr oder weniger starken dauernden Einschränkung oder Behinderung auswirken können.

Der Häufigkeit der Nennungen nach folgt dann die Kategorie "Prognose" (bei Med 1); gemeint ist eine schlechte Prognose. Von der inhaltlichen Bedeutung her gibt es enge Berührungen zu Therapierbarkeit und Lebensgefahr. Prognose wurde als eigene Kategorie belassen, da die Terminologie häufiger gebraucht wurde, auch sehr viel allgemeiner ist. Zu einem recht geringen Anteil werden als erstes Kriterium "Dauer, Chronizität" der Erkrankung und vom Beurteiler nachempfundene "seelische Belastung" genannt. Eine relativ große Anzahl von Nennungen vor allem bei Med 1 (15 %) war nicht zu kategorisieren und ist als "Sonstiges" aufgeführt. Die Nennung inhaltlich eher divergierender Gesichtspunkte ist bei Med 2 häufiger, vor allem taucht

hier das Kriterium "Beeinträchtigung der Persönlichkeit", bzw. Beeinträchtigung "der geistigen/intellektuellen Leistungsfähigkeit" auf, welches bei Med 1 noch gar nicht genannt wurde, von uns bei Med 2 unter der Gruppe Sonstiges aber subsumiert wurde.

Generell werden vorzüglich Kriterien genannt, die der konventionellen, eher von somatischen Aspekten her geprägten Beurteilung des Schweregrades von Krankheiten entsprechen. Erst in der Rangfolge an zweiter, mehr noch erst an dritter Stelle kommen Gesichtspunkte, welche spezifischer für psychiatrische Erkrankungen sind: seelische Belastung, vor allem auch "soziale Gesichtspunkte", worunter sowohl die soziale Belastung des Patienten selber als auch die Belastung der Angehörigen durch den Patienten verstanden wird. Erst bei der Nachuntersuchung, bei Med 2 also, wird ein ganz spezifischer Gesichtspunkt genannt, wie wir bereits erwähnt haben, nämlich "Verlust/Beeinträchtigung psychischer Funktionen".

Auch bei einer Zusammenstellung aller von beiden Gruppen überhaupt genannten Kriterien ohne Berücksichtigung der Rangfolge (Abb. 5, unterstes Histogramm) führen in beiden Gruppen die Nennungen von Therapierbarkeit, dann etwa gleich häufig andere Kriterien, bei Med 2 minimal häufiger für seelische Erkrankungen eher spezifische Gesichtspunkte wie Beeinträchtigung der Lebensqualität, seelische Belastung, soziale Gesichtspunkte.

Weiterhin haben insgesamt gesehen die überwiegend genannten Kriterien einen engen inhaltlichen Zusammenhang, so daß man fast mit gleichem Recht noch einige Kategorien in eine hätte zusammenfassen können, wie z. B. Therapierbarkeit mit Prognose und Lebensgefahr.

Es wird auch hieraus deutlich, daß das Beurteilungskriterium "schlimm" unter psychologischem Aspekt in

hohem Maße eindimensional ist. Ferner, was fast trivial erscheint und auch bei anderen empirischen Untersuchungen immer wieder gefunden wurde, daß eine als kaum therapierbar, lebensgefährlich und mit schlechter Prognose behaftete Krankheit wie Krebs eindeutig als schlimmste beurteilt wird. Dann aber folgt in der Skala eine schwere körperliche Krankheit, Querschnittslähmung, sogleich, da sie als Unfallfolge zwar nicht unbedingt lebensgefährlich, aber unheilbar ist und schwere körperliche Behinderung mit einschneidenden Veränderungen der Lebensführung verursacht. Mit gewissem Abstand folgt Schizophrenie, sicherlich wegen der schweren bis in relativ vielen Fällen überhaupt mangelnden Therapierbarkeit, der in vielen Fällen schlechten Prognose und der Beeinträchtigung des Lebens des Patienten. Dabei könnte die Veränderung bei Med 2 gegenüber Med 1, daß der Abstand zu Herzinfarkt größer wird, schon eine Auswirkung einer genaueren Kenntnis der Beeinträchtigungen sein, welche durch eine Erkrankung dieser Art verursacht werden, und einer etwas häufigeren auf den spezifischen Aspekt psychischer Erkrankungen eingehenden Beurteilung. Entsprechend kann man vermuten, daß durch das relative (aber geringfügige) Zurücktreten mehr somatischer Krankheiten, vor allem aber durch eine genauere Kenntnis aufgrund des Studiums, Hirnarteriosklerose relativ milder beurteilt wird. Es scheint, daß bei relativ uninformierten Beurteilern (vgl. dazu später die Gruppen Kontr 1, Tb, Alk) zuerst eine massive somatische Konnotation entsteht, verursacht durch den Wortbestandteil "Hirn".

Für die Medizinstudenten wurde nicht vorgegeben, beim Paarvergleich der Krankheiten von der Vorstellung auszugehen, daß sie selber betroffen sein könnten. Sie wurden vielmehr nach Bearbeitung der Paarvergleichsliste gebeten anzugeben, von welchem Gesichtspunkt her sie sich insgesamt entschieden hätten: eher aus der

Sicht des Arztes (A), aus der Sicht des Patienten, des Betroffenen (B), und da mehrere verlangten, auch ausdrücken zu können, daß das für sie schwierig zu sagen sei (was sehr realistisch erscheint), wurde als dritte Möglichkeit eingeräumt, "schwierig" (S) anzugeben.

Es ergab sich die Verteilung der Kennzeichnungen von Tab. 12.

Tabelle 12:
Gesichtspunkte der Entscheidung im Paarvergleich der Krankheiten bei den Beurteilergruppen Med 1 und Med 2 (absolute (f) und relative (f%) Häufigkeiten)

	Med 1		Med 2	
	f	f%	f	f%
A(rzt)	19	28.8	6	16.2
B(troffener)	38	57.6	19	51.4
S(chwierig)	6	9.1	6	16.2
keine Angabe	3	4.5	6	16.2
	66	1oo.o	37	1oo.o

In beiden Gruppen haben die meisten angehenden Ärzte zumindest die Intention, sich in einen Patienten, in den Betroffenen, hineinzuversetzen. Vielleicht sind jedoch die in der Minderheit befindlichen Beurteiler, welche das für einen (angehenden) medizinischen Fachmann für schwierig halten, realistischer. Die Kriterien der Beurteilung, die vorher besprochen wurden, deuten in ihrer vorherrschenden Gewichtung eher auf medizin-immanente, eher die Krankheit distanziert betrachtende Gesichtspunkte hin, vor allem auch in ihrer Formulierung: Therapierbarkeit, Lebensgefahr (oft als "letaler Ausgang" benannt). Diese Kriterien sind sicher nicht inadäquat, mildern aber evtl. die Einschätzung von solchen Krankheiten relativ stärker, die üblicherweise keine starke somatische Gefährdung oder Beein-

trächtigung mit sich bringen.

4.3 Die Patientengruppen als Beurteiler

Das Interesse an der Einschätzung psychiatrischer Erkrankungen legte nahe, eine Gruppe von selbst betroffenen Patienten auszuwählen. Es bot sich eine Gruppe von Alkoholkranken in der Behandlung einer Fachklinik an, einmal aufgrund der Erreichbarkeit, zum anderen deswegen, weil eine Akzeptanz der Diagnose hier eine notwendige Voraussetzung für eine solche Behandlung ist, für die Befragungsmethode aber ebenso notwendig. Damit ist gewährleistet, daß die Beurteiler ihr Krankheitsbild wiederfinden. Gleiche Gründe sprachen für die Auswahl von Tuberkulosepatienten als Vertreter einer Gruppe von organisch Erkrankten, wobei sich herausstellte, daß die Erreichbarkeit weniger gut gegeben ist als bei Alkoholpatienten. Es war aber auch hier mit Sicherheit vorauszusetzen, daß persönliche Klarheit über Art und Bezeichnung des eigenen Krankheitsbildes besteht, bedingt durch die Aufnahme ein eine entsprechende Spezialbehandlung.

Wie stufen selbst betroffene Patienten ihre Erkrankung in Relation zu anderen Erkrankungen ein? Einerseits könnte man erwarten, daß die eigene Erfahrung mit den Belastungen der Behandlung und den Beeinträchtigungen, welche folgen, die Betroffenen ihre Krankheit wesentlich schlimmer einstufen läßt als Nichtbetroffene es täten. Andererseits könnte auch angenommen werden, daß durch eigene Erfahrung die Krankheit sozusagen "entmythisiert" wird, daß sie realer und nüchterner und evtl. willentlich forciert leichter genommen wird als von anderen Gruppen.

In unseren Gruppen Tb und Alk bestätigt sich keine der beiden Vermutungen (Tab. 13 und 14; Abb. 1). Bis auf eine Gruppe leichter Krankheiten, Grippe, Akne, Blinddarmentzündung, bei denen es Vertauschungen der Skalen-

Tabelle 13:

Skalenwerte der Gruppe Tb

N = 48

Coefficient of Agreement U = o.452

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.48o	o.ooo	o.174
Akne	-o.oo8	1.19o	o.284	o.1o3
Neur	o.855	o.9o5	1.619	1.113
Blin	-o.o81	1.1o6	o.332	o.ooo
Blut	o.569	o.82o	1.299	o.827
Alko	1.357	o.795	2.255	1.594
Zuck	1.o4o	o.429	1.877	1.298
Syph	1.179	o.749	2.o5o	1.444
Hirn	1.58o	o.7o3	2.488	1.838
Tb	1.336	o.831	2.232	1.67o
Herz	1.597	1.178	2.6o9	2.176
Depr	1.o17	o.6o2	1.811	1.27o
Quer	2.135	1.223	3.537	2.484
Krbs	2.526	1.724	4.771	2.883
Schi	1.5o5	1.264	2.596	1.763
Chi^2:	224.552		2o3.445	151.44o
S_D :				3.156
S_L :				65.775
S_T :				76.259
Rho :				.8562
Rss :				.9522

Korrelationen (r)

Skala V zu Skala III: .987

Iteration zu Skala V : .995

Iteration zu Skala III: .978

Tabelle 14:

Skalenwerte der Gruppe Alk

N = 41

Coefficient of Agreement U = o.487

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.465	o.ooo	o.ooo
Akne	o.3o2	1.36o	o.5o9	o.66o
Neur	1.o46	o.5oo	1.8o4	1.486
Blin	o.28o	o.894	o.885	o.454
Blut	o.454	o.746	1.o79	o.893
Alko	1.341	o.827	2.136	1.781
Zuck	o.995	o.591	1.732	1.435
Syph	1.366	o.675	2.166	1.8o6
Hirn	1.87o	1.532	3.128	2.537
Tb	1.313	o.57o	2.o32	1.842
Herz	1.83o	o.876	2.678	2.497
Depr	1.431	o.685	2.168	1.992
Quer	2.o8o	1.783	3.665	2.852
Krbs	2.471	1.349	4.oo6	3.o32
Schi	1.826	1.147	2.9o6	2.266
Chi^2 :	299.1o3		296.29o	154.699
S_D :				3.775
S_L :				71.181
S_T :				84.169
Rho :				.8381
Rss :				.9477

Korrelationen (r):

Skala V zu Skala III: .987

Iteration zu Skala V : .993

Iteration zu Skala III: .983

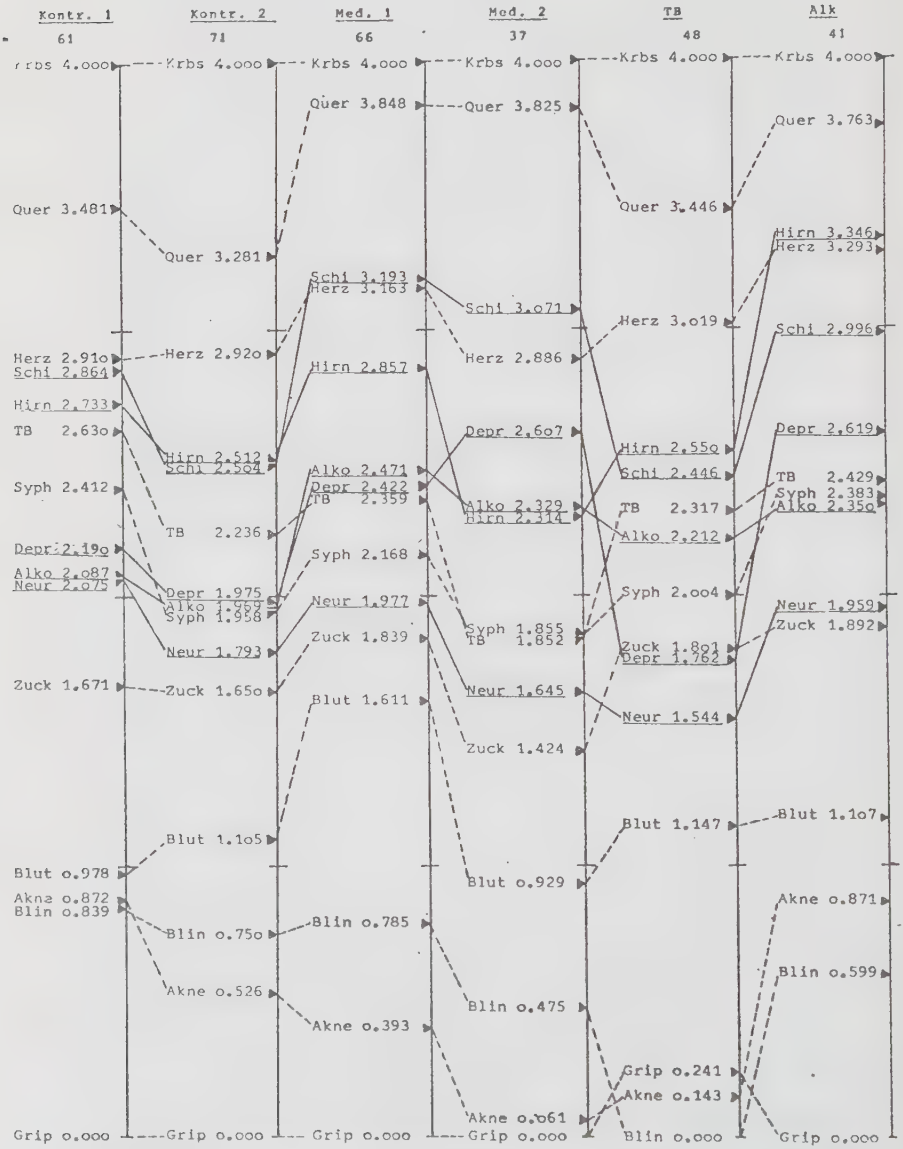
plätze in Relation der beiden Gruppen zueinander wie zu den beiden Skalen Med 1 und Med 2 gibt, halten sie für die körperlichen Krankheiten die gleiche Reihenfolge ein wie die Medizinergruppen. Die Gruppe Tb ordnet Tuberkulose den mittelschweren Erkrankungen zu; in Relation zu der bei dieser Gruppe am geringsten gespreizten Skala geringfügig über dem Mittelpunkt der Skala (vom Maximum aus gesehen), was aber auch bei den beiden von Med 1 und Alk erstellten Skalen zutrifft. Med 2 stufte Tuberkulose geringfügig unter dem Mittelpunkt der Skala ein, vertauscht gegenüber allen anderen Gruppen (Abb. 1) die Position von Syphilis und Tuberkulose, bei allerdings praktisch identischen Skalenwerten.

Ähnlich beurteilt die Gruppe Alk Alkoholismus weder deutlich schwerer noch deutlich leichter in Relation zu den anderen Gruppen. Auch sie ordnet diese eigene Krankheit im Mittelbereich ein, leicht über dem arithmetischen Mittel der eigenen Skala.

Da durch die unterschiedliche Spreizung der Skalen bei den einzelnen Beurteilergruppen ein Vergleich relativer Positionen von Krankheiten zueinander oder zum Mittel der eigenen Skala erschwert wird, findet sich in Abb. 6 eine graphische Darstellung der Skalen nach einer Skalentransformation, welche die verschiedene Spreizung der Skalen bei den einzelnen Gruppen eliminiert. Die Transformation wurde so durchgeführt, daß jeweils - wie bisher - die leichteste Krankheit, zumeist Grippe, als Nullpunkt gesetzt wird, die als schwerste beurteilte Krankheit, in allen Fällen Krebs, mit 4.o angesetzt wird. Entsprechend sind dann jeweils die anderen Werte linear transformiert (mit einem pro Beurteilerskala konstanten Faktor multipliziert).

Unterschiede der Skalen der Patientengruppen zu den Skalen der Medizinergruppen zeigen sich in der Bewertung der psychiatrischen Krankheitsbilder. Während

Abbildung 6:
Skalen aller Beurteilergruppen nach linearer Skalentransformation, so daß unterschiedliche Spreizung eliminiert wird



(lin. Transformation für jede Gruppe derart, daß jeweils ein maximaler Skalenwert von 4.0 resultiert)

Med 1 und stärker dann noch Med 2 Schizophrenie als schwerste psychische Erkrankung ansehen, setzt die Gruppe Alk Hirnarteriosklerose (Verkalkung) an die erste Stelle der psychiatrischen Krankheitsbilder (Abb. 1 und Abb. 6), in praktisch identischer Position mit Herzinfarkt. Auch die Gruppe Tb setzt Hirnarteriosklerose numerisch an die oberste Position der psychiatrischen Erkrankungen, jedoch niedriger als Herzinfarkt und auf faktisch die gleiche Position wie Schizophrenie. Und obgleich die Gruppe Alk ähnlich wie Tb eine in Relation zu den anderen Gruppen weniger gespreizte Skala aufweist, also nach unserer Interpretation insgesamt weniger zwischen den Krankheitsbildern differenziert, unterscheidet diese Gruppe im Gegensatz zu Tb die psychiatrischen Erkrankungen Hirnarteriosklerose, Schizophrenie, endogene Depression, Alkoholismus und Neurose nach der Reihenfolge der Schwere deutlicher voneinander als die Gruppe Tb. Bei diesen gehören nach der Reihenfolge der Schwere Hirnarteriosklerose, Schizophrenie, Alkoholismus, Depression, Neurose zu einer auf der Skala dicht beieinander liegenden Mittelgruppe von Krankheiten, welche noch Tuberkulose, Syphilis und Zuckerkrankheit umfaßt. Nach oben abgesetzt (Abb. 1 und Abb. 6) sind bei der Gruppe Tb Krebs, Querschnittslähmung und Herzinfarkt, nach unten Bluthochdruck, wiederum auf fast gleicher Position nahe am Nullpunkt Grippe, Akne und Blinddarmentzündung.

4.4 Skalierung der Krankheiten und Informiertheit bei den Patientengruppen

Die erste Intention für die Aufnahme eines "Wissenstests" über die zu skalierenden Krankheiten in die Untersuchung war, ein Lerninstrument zu schaffen. Zugleich sollte aber auch versucht werden, wenn auch sehr grob aufgrund der angestrebten und auch durchgängig erreichten geringen Schwierigkeit des Tests (vgl. Anhang 3), eine gewisse Kontrolle des Infor-

miertheitsgrades bei den Gruppen der Nicht-Mediziner zu erhalten. Die Kennwerte für die Schwierigkeit der Items, damit zugleich für die Lösungshäufigkeiten für die einzelnen Items finden sich für alle Items des Wissenstests im Anhang 3, wo sie jeweils dem Wortlaut der Items der besseren Übersichtlichkeit halber vorausgestellt sind.

Tab. 15 soll einen Vergleich der relativen Kenntnis einzelner Krankheitsbilder erleichtern. Für jede Beurteilergruppe und jedes Krankheitsbild sind die sich aus der Häufigkeit richtiger Beantwortung der drei jeweils zutreffenden Items für jede Krankheit ergebenden mittleren Prozentzahlen angegeben (im Bewußtsein, daß die Mittelung von Prozentwerten, dann zumal von nur drei Werten, nicht "lege artis" ist).

Die Tabelle zeigt in den beiden letzten Spalten hinreichend deutlich, daß der Test für die Gruppen Tb und Alk durchaus leicht ist, daß aber die Gruppe Alk ihn doch besser bewältigt als Tb. Bei Alk liegen bis auf zwei Krankheiten alle mittleren Lösungshäufigkeiten für die einzelnen Krankheiten über 9o %. Bei Tb gibt es keine Lösungshäufigkeit unter 8o %, aber nur bei 7, also nicht ganz der Hälfte von 15 Krankheitsbildern, eine über 9o %.

Die in Klammern angeführten Rangplätze sollen einen ebenso groben, aber schnellen Überblick über den relativen Bekanntheitsgrad der einzelnen Krankheiten pro Beurteilergruppe geben. Dabei ist allerdings zu bedenken, daß die Maßzahl "Schwierigkeitsindex" wesentlich von zwei Faktoren bestimmt wird: Einmal von der Kenntnis der Probanden über das erfragte Objekt, zum anderen davon, wie schwierig ein Statement dann in Relation zum Objekt (formuliert) ist. Um das näher zu erläutern: Die Items zu Grippe sind offensichtlich nicht glücklich formuliert. Die Schwierigkeitsindices sind für den hier abgedeckten Bereich relativ niedrig,

Tabelle 15:
"Wissenstest" über die Krankheiten - gemittelte Lösungshäufigkeit (Schwierigkeitsindex) der drei zutreffenden Items und daraus sich ergebende Rangordnung (Rangplätze in Klammern) für die "Kenntnis" der Krankheiten

	Untersuchungsgruppen			
	Kontr 1	Kontr 2	Tb	Alk
Akne	92.6 (6)	95.6 (13)	89.o (8)	95.3 (9)
Blut	89.6 (2)	9o.3 (7)	92.o (9)	93.o (6)
Alko	94.o (8)	93.3 (1o)	87.7 (7)	97.7 (13)
Herz	9o.7 (4)	93.o (9)	86.3 (6)	9o.3 (3)
Schi	94.3 (9.5)	89.7 (6)	85.o (4)	92.o (5)
Blin	96.o (12)	91.7 (8)	97.3 (15)	95.3 (9)
Krbs	99.3 (15)	99.o (15)	95.3 (13)	95.3 (9)
Neur	94.7 (11)	85.3 (2)	82.7 (2.5)	87.7 (2)
Grip	85.7 (1)	79.3 (1)	85.7 (5)	86.3 (1)
Hirn	9o.o (3)	89.3 (5)	82.7 (2.5)	91.3 (4)
Zuck	97.7 (13)	94.3 (11)	93.3 (1o.5)	98.7 (14.5)
Depr	91.7 (5)	86.3 (4)	8o.7 (1)	94.7 (7)
Syph	93.7 (7)	95.o (12)	93.3 (1o.5)	96.7 (12)
Quer	98.7 (14)	96.7 (14)	96.7 (14)	98.7 (14.5)
Tb	94.3 (9.5)	86.o (3)	93.7 (12)	96.o (11)

was heißt, es sind die schwierigsten Items. Daraus ist aber sicher nicht mit Erstaunen zu folgern, daß Grippe die am wenigsten bekannte Krankheit sei, sondern in diesem Falle eben, daß die Items zu schwierig gewählt bzw. zu schwierig formuliert wurden.

Diese gegenseitige Abhängigkeit wird bekanntlich erst durch eine formelle Konstruktion und Standardisierung eines Tests ausgeschaltet, worauf hier aus guten Gründen verzichtet wurde.

Die Beziehung zwischen genereller Informiertheit über die Krankheiten und Skalierung nach dem Grad subjektiv empfundener Schwere wird aber vor allem bei der Gruppe Tb aus dem Vergleich der Kenngrößen deutlich. Die Skala der Gruppe zeigt das relativ geringste Maß von Differenzierung, die Kennzahlen für den Wissenstest zeigen das relativ (relativ zu den anderen Gruppen) geringste Maß von Kenntnis. Es wurde bereits herausgestellt, daß es mehrere Faktoren gibt, welche die Spreizung einer Skala im Paarvergleich beeinflussen; hier zeigt sich, daß auch die Kenntnis von Krankheiten einer der bestimmenden Faktoren ist.

Um den Überblick über das Ergebnis des "Wissenstests" zu vereinfachen, wurde für jeden Beurteiler ein Gesamtrohwert (GRW) errechnet. Da die Beantwortung der jeweils 5 Einzelitems für jede Krankheit aufgrund der Instruktion und des Inhalts der Statements nicht stochastisch unabhängig war (was beabsichtigt war), erschien folgende Auswertungsregel adäquat: Für jedes angekreuzte Item beim jeweiligen Krankheitsbild erhielt der Beurteiler den Wert 1 zugeschrieben, für jedes nicht angekreuzte den Wert o. Dann wurden die Werte der richtigen Items pro Krankheit addiert, ebenso die Werte der falschen Items, die Summe der falschen von der Summe der richtigen subtrahiert und dann (da die Instruktion bei Befolgung auf jeden Fall einen Treffer garantierte) noch einmal vom Ergebnis 1

subtrahiert. Formelhaft lautet also die Rechenregel:

Rohwert Krankheit = (R + R + R) - (F + F) - 1.

Es kam vor, daß Beurteiler weniger als 3 Items pro Krankheit entgegen der Instruktion ankreuzten. So konnte sich ein Rohwert pro Krankheitsbild von -3 bis maximal +2 ergeben. Zur Bestimmung des Gesamtrohwertes (GRW) als Kennwert für die Leistung im sogenannten Wissenstest wurden über die 15 Krankheiten hinweg nur die positiven Werte addiert, die negativen außer acht gelassen. Die erreichten Gesamtrohwerte (GRW) der Gruppen, denen der Wissenstest vorgegeben wurde (Tb, Alk, Kontr 1 und Kontr 2) finden sich als Häufigkeitstabelle in Tab. 16. Für jede Gruppe stehen jeweils in der ersten Spalte die absoluten, in der zweiten Spalte die relativen Häufigkeiten. Bei den relativen Häufigkeiten ist die Lage des Medians (Md) wie die Grenze des ersten Quartils (Q1) gekennzeichnet.

Wie zu sehen ist, zeigen die Verteilungen die dem angestrebten geringen Schwierigkeitsgrad entsprechende rechtsschiefe Form.

Die Mediane der vier Gruppen differieren zum Teil. Ebenso sind die Dispersionen (Interquartilstreuung) verschieden. Eine statistische Prüfmöglichkeit auf Unterschiede bietet der H-Test (SIEGEL 1956, LIENERT 1973), wenn er auch eigentlich nur Unterschiede in der zentralen Tendenz (hier: Median) erfassen soll und seine Beeinflußbarkeit durch unterschiedliche Dispersionen üblicherweise unerwünscht ist. Für die hier geforderte überblickshafte Prüfung erscheint sein Ansprechen auf unterschiedliche Dispersionen (LIENERT 1973) jedoch sogar willkommen. Die Prüfung über alle vier Gruppen hinweg erbringt einen sehr signifikanten Wert für H (korrigiert für gleiche Rangplätze) von 11.4 ($p < 0.01$). Bei Prüfung zwischen den einzelnen Gruppen mit dem Verfahren von NEMENYI (DIEHL & KOHR 1977, LIENERT 1973) zeigt jedoch keine

Tabelle 16:
Ergebnis des "Wissenstests" - absolute (f) und relative (f%) Häufigkeiten eines erreichten Gesamtrohwertes (GRW) in den einzelnen Gruppen

	Gruppen							
	Kontr 1		Kontr 2		Tb		Alk	
GRW	f	f%	f	f%	f	f%	f	f%
3			1	1				
4					1	2		
5					1	2		
6								
7					1	2		
8					3	6		
9					1	2		
1o			4	6	1	2		
11							1	2
12			1	1			1	2
13							1	2
14	3	5			1	2	1	2
15	1	2			1	2		
16			5	7	2	4	1	2
17	1	2						
18	1	2	4	6	1	Q1:2	2	5
19					1	2		
2o	1	1o	3	4	5	1o	2	5
21				Q1:-				
22	2	3	8	11	2	4	2	Q1:5
23							1	2
24	12	Q1:2o	15	Md21	8	Md17	4	1o
25	3	5	1	1				
26	9	Md15	9	13	5	1o	8	Md2o
27	1	2					1	2
28	1o	16	14	2o	12	25	13	32
29	1	2						
3o	11	18	6	8	2	4	3	7
N =	61		71		48		41	
Median GRW =	25.67		24.13		23.88		26.o6	

(Kennzeichnung des Medians: Md, der Grenze des 1. Quartils als Streuungsmaß: Q1)

Prüfgröße Signifikanz (α = o.o5).

Nach Inspektion der Verteilung dürfte daher diese Signifikanz nicht auf die Unterschiede der Mediane, sondern vor allem auf die unterschiedlichen Streuungen zurückzuführen sein, und hier insbesondere auf die gerade interessierende Gruppe Tb, bei der sich die GRW fast über das ganze Spektrum verteilen. Das ist als weiterer Beleg dafür anzusehen, daß der Kenntnisstand in einer Gruppe die Spreizung der Skala und damit die Differenzierung zwischen den zu beurteilenden Objekten beeinflußt.

Zur Kennzeichnung der Beurteilergruppen sind Kennwerte für Schul- und Berufsausbildung erhoben worden, um einen Indikator für den allgemeinen Bildungsstand und damit indirekt für die allgemeine intellektuelle Leistungsfähigkeit zu erhalten.

Von Interesse ist, ob ein Zusammenhang zwischen Kenntnis über die Krankheiten (GRW), allgemeinem Bildungsstand und Anzahl der zirkulären Triaden (ZT) im Paarvergleich besteht. Es fehlt für diese Prüfung noch eine sinnvolle Kenngröße für den allgemeinen Bildungsstand (BIST).

Es ist prinzipiell möglich, schulische und berufliche Ausbildung einer Rangfolge zuzuordnen. So wurde aus den Angaben der Beurteiler zu ihrer Schul- und Berufsausbildung ein kombinierter Score zur Kennzeichnung des Bildungsstandes (BIST) gebildet, um durch eine größere Variationsbreite dieser Kenngröße besseren Aufschluß zu erhalten über Zusammenhänge zwischen Bildungsniveau, Anzahl intransitiver Urteile im Paarvergleich und Informiertheitsgrad über Krankheiten (GRW des "Wissenstests"). Dabei ist zu beachten, daß die Kennwerte für Zusammenhänge (wir haben uns auf Rangkorrelationen beschränkt) aufgrund der trotz prinzipieller Legitimität etwas gewaltsamen und vereinfachenden Bildung der Variablen mit größter Vorsicht zu

interpretieren sind.

Der Kennwert für den Bildungsstand eines Beurteilers ergab sich durch Zusammenfassung eines Punktsystems für Schul- und Berufsausbildung. Den Angaben über die Schulbildung (vgl. das Fragebogenformular Anhang 2) wurden folgende Punktwerte zugeordnet:
ohne Abschluß = 1, Hauptschule und Sonstiges = 5, Realschule/Fachschule = 9, Abitur = 14; für die Berufsausbildung: keine = o, angelernt sowie noch in Ausbildung = 1, Lehrberuf = 3, Studium = 4. Bei additiver Kombination dieser Punktwerte ergibt sich eine Rangskala, bei welcher die Schulbildung die Basis der Einordnung darstellt, die hinzukommende Berufsausbildung diese Basisgruppierung dann weiter differenziert.

Die drei obengenannten Kenngrößen allgemeiner Bildungsstand (BIST), Gesamtrohwert im "Wissenstest" (GRW) und Anzahl intransitiver Urteile, d. h. zirkulärer Triaden im Paarvergleich (ZT) sind in Tab. 17 über den Eta-Koeffizienten (MITTENECKER 197o, FRÖHLICH & BECKER 1971; vgl. LIENERT 1973 und 1978) in Beziehung gesetzt. Andere Meßgrößen für den Zusammenhang kommen aufgrund der differierenden Verteilungsformen zumal in den Gruppen Tb und Alk nicht in Betracht. Der an sich vorzeichenlose Koeffizient bezeichnet analog zu einem Korrelationskoeffizienten die Stärke eines Zusammenhangs in den Grenzen o.oo bis 1.oo, wobei (aus der einfachen Varianzanalyse abgeleitet) die in der Tabelle an erster Stelle stehende Variable jeweils als unabhängige, hier wohl besser "vorgängige" genommen wird, nach deren Stufungen die Veränderung der nachfolgenden betrachtet wird (weswegen bei Umkehrung nicht wie beim Korrelationskoeffizienten r oder R_s der gleiche Wert resultieren muß). Bei Signifikanz kann das Vorzeichen, das die Richtung des Zusammenhangs kennzeichnet, nach Inspektion der Daten ergänzt werden.

Die Erwartung war, daß sich zwischen allen drei Variab-

Tabelle 17:

Zusammenhang zwischen der kombinierten Variable Bildungsstand (BIST), Anzahl der zirkulären Triaden im Paarvergleich (ZT) und Gesamtrohwert im "Wissenstest" (GRW) (Eta-Koeffizient)

	Tb	Alk	Kontr 1	Kontr 2
GRW zu ZT	(-).69 s $F_{38;9}$=3.89	.30 ns $F_{34;6}$=0.56	(-).62 ss $F_{53;7}$=4.75	(-).53 s $F_{62;8}$=3.06
BIST zu GRW	.37 ns $F_{42;5}$=1.34	.46 ns $F_{34;6}$=1.53	.49 ns $F_{54;6}$=2.84	(+).59 s $F_{63;7}$=4.87
BIST zu ZT	.44 ns $F_{42;5}$=2.03	.33 ns $F_{34;6}$=0.67	.30 ns $F_{54;6}$=0.89	.43 ns $F_{63;7}$=2.02

len überzufällige Zusammenhänge zeigen. Bestätigt wird diese Erwartung nur bei drei von vier Untersuchungsgruppen für den Zusammenhang zwischen Kenntnis über Krankheiten und Konsistenz des individuellen Urteilsverhaltens im Paarvergleich (GRW zu ZT). Weiterhin sind diese Zusammenhänge nicht sehr stark ausgeprägt; nur der jeweils kleinere Teil der gesamten Varianz beider Variablen ist gemeinsam (Quadrat des Eta-Koeffizienten, sog. "Korrelationsverhältnis": bei Tb 48 %, bei Kontr 1 38 %, bei Kontr 2 28 % gemeinsame Varianz).

Für die Variablen Bildungsstand (Schul- und Berufsausbildung; BIST) und Kenntnis der vorgestellten Krankheitsbilder, erfaßt im "Wissenstest" (GRW), zeigt sich nur für Kontr 2 ein überzufälliger Zusammenhang (Eta = (+).59, Eta^2 = .35, d. h. 35 % gemeinsame Varianz), der auch nur mäßig stark ausgeprägt ist. Ein überzufälliger Zusammenhang zwischen BIST und ZT läßt sich in keiner der Gruppen feststellen.

Zur Erklärung dieses Ergebnisses ist mit Sicherheit in erster Linie ein statistischer Grund heranzuziehen: Es dürfte sich hier die (durchaus mangelnde) Größe der Stichproben über eine daraus resultierende nichtnormale Verteilungsform der Variablen auswirken, bei der Variablen GRW wohl auch dadurch bedingt, daß aufgrund der Doppelfunktion (Wissenstest und Informationsvermittlung) eine nicht-normale Verteilung bei angezielter geringer Schwierigkeit bewußt in Kauf genommen wurde.

Von der inhaltlichen Seite her erscheint folgende Interpretation angemessen: Für die Konsistenz des Urteilsverhaltens im Paarvergleich ist die Kenntnis der Krankheiten ein bestimmender Faktor, aber wie bereits erörtert, nur einer unter mehreren. Gering ausgeprägte Unterschiede in der subjektiv empfundenen Schwere dürften für das Auftreten intransitiver Urteile von deutlich größerer Bedeutung sein, während eine mög-

liche Mehrdimensionalität, wie durch die Indices der (linearen) Skalierbarkeit ausgewiesen, eher zurücktritt. Das Bildungsniveau wirkt sich in den vier Gruppen nur einmal meßbar aus, und zwar bei Kontr 2 über den Zusammenhang zum Ergebnis des "Wissenstests". Der Zusammenhang zwischen Bildungsniveau und Kenntnis über Krankheiten, der zu postulieren ist, fällt gewissermaßen dem Untersuchungsdesign zum Opfer. Daß dieser Zusammenhang besteht, erhellt daraus, daß ein großer Anteil der ursprünglichen Gruppe Tb, welche durch das geringste Bildungsniveau der vier Gruppen charakterisiert wird, an der Bearbeitung der Paarvergleichsprozedur scheiterte und deswegen ausgeschieden werden mußte (22 von ursprünglich 7o; vgl. Tab. 5).

4.5 Kriterien der Patientengruppen für die Einstufung der Krankheiten

Die Gruppen Med 1 und Med 2 waren nach der Bearbeitung der Paarvergleichsliste gebeten worden, die Kriterien ihrer Beurteilung in Stichworten zu notieren und zusätzlich anzugeben, ob sie sich eher in die Lage des behandelnden Arztes oder des Betroffenen hineinversetzt hätten.

Für die Patientengruppe wie auch die Kontrollgruppen entfiel die letzte Frage, da sie ausdrücklich gebeten wurden - bzw. daran erinnert wurden - sich in die Lage eines Betroffenen hineinzuversetzen.

Eine zusätzliche schriftliche Befragung nach den Kriterien war bei diesen Gruppen nicht möglich. Daher wurde bei allen einzelnen Untersuchungsdurchgängen (je einer für die beiden Kontrollgruppen; zwei Einzelgruppen für die Gruppe Alk, insgesamt sechs für die Gruppe Tb) jeweils mit einem freiwillig verbleibenden interessierten Teil der Gruppe ein Nachgespräch geführt. Diese Probanden waren auch daran in-

teressiert, etwas über den Gesamtzusammenhang der Untersuchung zu erfahren oder noch Fragen zu den einzelnen Krankheiten zu stellen. Sie konnten vom Untersucher global zu ihren Kriterien für die Beurteilung befragt werden sowie (bei den Patientengruppen) zu ihren Erfahrungen mit der eigenen Erkrankung.

Bei der Gruppe Alk zeigte sich, daß alle an einem Nachgespräch interessiert waren. Auch wurden kaum Verständnisfragen zu den einzelnen Krankheiten gestellt, vor allem fast keine Fragen zu den psychiatrischen Krankheitsbildern. Es schien so, und das Ergebnis der "Wissenstests" stützt diese Annahme, daß in dieser Gruppe eine insgesamt klarere Vorstellung von diesen Erkrankungen vorhanden war als in der Gruppe Tb.

Was die Kriterien der Beurteilung angeht, zeigte sich, daß entgegen der Erwartung spontan fast gleiche Gesichtspunkte genannt wurden wie von der Gruppe der Medizinstudenten. Dabei war eine exakte Gewichtung aufgrund der ungeordneten Nennungen im Gruppengespräch allerdings nicht möglich. Genannt wurden hier als Kriterien vor allem, wie weit eine Krankheit heilbar sei, ob sie lebensgefährlich sei, ob sie schwere gesundheitliche (d. h. körperliche) Schäden hinterlasse. Psychische Behinderung, emotionale Belastung oder soziale Aspekte wurden von der Gruppe spontan nicht genannt.

Die eigene Erkrankung sah man unter den Aspekten, die auch in einer klassischen Langzeittherapie des Alkoholismus betont werden: Daß es das schwerste Problem gewesen sei einzusehen, daß man wirklich krank sei, daß man seine Sucht nicht beherrschen konnte, daß man sich ferner in der Folge darauf einzustellen habe, völlig "trocken" zu bleiben. Die Bedeutung des Alkoholismus für die eigene Stellung im sozialen Leben, in der Familie und im Beruf oder etwa die Bela-

stungen für die Familie oder Partnerschaft wurden von der Gruppe selbst nicht angesprochen.

Es fiel auf, daß die Äußerungen kaum retrospektiv waren, daß man vor allem darauf zu sprechen kam, was man in Zukunft tun müsse; dabei war die Grundhaltung eher optimistisch.

Hier akzentuierten die verschiedenen Untergruppen der Tuberkulose-Patienten anders. Zum einen stellten sie sehr viel mehr Fragen zu einzelnen Krankheiten, vor allem die jüngeren Patienten. Gefragt wurde - teilweise schon vorher (vgl. Erläuterungen zu Nachfragen bei der Instruktion zum Paarvergleich) - bei den körperlichen Erkrankungen nach Akne, dann interessierten sich wiederum vor allem die Jüngeren sehr für psychiatrische Erkrankungen, hier vor allem für Neurose und für den Unterschied zwischen endogener und reaktiver Depression, wollten auch wissen, wie gut man so etwas behandeln könne.

In gleicher Weise wie die Alkoholpatienten nannte aber auch diese Gruppe in etwa die gleichen Kriterien für die Einordnung der Krankheit: Heilbarkeit/Behandelbarkeit (wie sie zu den einzelnen Krankheiten wohl vermutet oder gewußt wurde), Lebensgefahr. Soziale Folgen und emotionale Belastungen oder eine Beeinträchtigung der psychischen Funktionen wurden von den Tuberkulose-Patienten selbst nicht genannt.

Dieser letzte Aspekt kam nur in den Berichten über das Erleben der eigenen Erkrankung zur Sprache, und nur dann, wenn vom Untersucher direkt danach gefragt wurde. Ein wesentliches Moment war nach dem Bericht einiger Patienten die anfängliche Unsicherheit und Angst, was für eine Erkrankung tatsächlich vorläge. Dabei stand im Vordergrund die Angst, die Erkrankung könne sich vielleicht als Lungenkrebs herausstellen. Insofern wurde die endgültige und sichere Diagnose einer Tuberkulose als Erleichterung gesehen, von da-

her die eigene Krankheit auch nicht als besonders schwer bezeichnet, da sie heute mit Sicherheit heilbar sei. Diese Angstthematik wurde vor allem bei solchen Patienten angesprochen, welche in Fachkliniken behandelt wurden, die neben der Tuberkulose-Behandlung (oder auch hauptsächlich) andere Lungenerkrankungen, zu einem großen Teil eben bösartige Lungenerkrankungen behandelten.

Ein zweites für die Gruppe Tb bezeichnendes Moment war abhängig von den offensichtlich unterschiedlichen Behandlungsstilen, die Patienten in einer Erstbehandlung erfahren hatten. Es war für einige das Erleben einer plötzlichen und radikalen Isolierung. Sehr eindrucksvoll berichtete eine Krankenschwester aus einer Universitätsklinik, die jetzt selber Patientin war, wie sie erlebt hatte, daß plötzlich alle Leute, zum großen Teil Bekannte, von ihr zurückwichen und eine breite Gasse öffneten, als sie jetzt als Patientin mit dem in jener erstbehandelnden Klinik üblichen Mundschutz die Vorhalle vor einem Hörsaal passieren mußte. Ähnlich schilderte ein Patient, wie das Gesundheitsamt die Wohnung desinfizierte, unter Beobachtung aller Nachbarn, und wie er von einem auf den anderen Tag von seiner Familie isoliert wurde.

Etliche Patienten erinnerten sich daran, daß die teilweise durchgeführte Isolierung sehr belastend war, vor allem aber die in jedem Fall erfolgte Trennung von den Angehörigen und der normalen Alltagswelt, dann das Verurteiltsein zum Nichtstun, obwohl man subjektiv keine Beschwerden spüre. Einige betrachteten auch die Unterbrechung einer beruflichen Ausbildung oder Karriere als schwer zu verkraften.

Ein weiterer Kontrast zwischen den beiden Patientengruppen ist bemerkenswert: Bei den Tuberkulose-Patienten (diejenigen, welche sich an den Gesprächen beteiligten) ergab sich der Eindruck, daß sie über Details

ihrer Erkrankung, über die medikamentöse Behandlung und darüber, wie es in Zukunft weitergehen solle, deutlich weniger orientiert und unsicherer waren als die Alkoholpatienten. Zum Teil wurde mangelnde Möglichkeit eines ausführlichen Gesprächs mit den behandelnden Ärzten, mangelnde Gelegenheit, ausgiebig Fragen zu stellen, sehr beklagt. Einige Patienten brachten die geschilderten Schwierigkeiten in Zusammenhang damit, daß sie bei zahlreichen Mitpatienten eine erhebliche Alkoholproblematik vermuteten, die entweder während der Behandlung verstärkt oder gar erst entstanden sei. Einzig für eine Teilgruppe in einer Klinik, die vorzugsweise Nachkuren durchführt, traf dieser Eindruck nicht zu. Hier wurden aber als indirekte Bestätigung gerade die guten Gesprächs- und Informationsmöglichkeiten im Gegensatz zu früheren Behandlungen hervorgehoben.

Es soll noch mit allen Einschränkungen der klinische Eindruck des Untersuchers festgehalten werden, daß ein größerer Teil der beteiligten Tuberkulose-Patienten ein psychisch deutlich beeinträchtigteres Bild bot als etwa die Gruppe der Alkoholpatienten, wo das eher erwartet worden war. Dieser Eindruck beruht nicht darauf, daß die Gruppe der Tuberkulose-Patienten im Mittel deutlich älter war. Zu vermuten sind Faktoren, die möglicherweise im Sinne einer negativen Selektion in Bezug auf das Erkrankungsrisiko wirksam sind, vielleicht aber auch auf die Zuweisung zu bestimmten Behandlungsinstitutionen, wohl gerade zu denen, die wir erreichen konnten. Dieser Eindruck paßt dazu, daß von insgesamt 7o ursprünglich beteiligten Patienten trotz ausführlicher und oft für den einzelnen wiederholter Instruktion (abgesehen von Patienten anderer Muttersprache, die zwar mitarbeiten durften, aber deren Antwortbogen von vornherein ausgeschieden wurden) 22 Patienten nicht in der Lage waren, die Paarvergleichsliste zu bearbeiten.

4.6 Ergebnisse der Kontrollgruppen

In der bisherigen Erörterung der Ergebnisse sind bereits öfter Ergebnisse der beiden Kontrollgruppen erwähnt worden bzw. zu einzelnen Vergleichen herangezogen worden.

Es sei noch einmal kurz auf die Charakterisierung der beiden Kontrollgruppen mit 61 bzw. 71 männlichen Beurteilern mit einem medianen Alter von etwa 2o Jahren bei geringer Altersstreuung hinzuweisen. Sie unterschieden sich auch in den sonstigen Kenngrößen nicht voneinander (s. o. Tab. 8). Der einzige Unterschied ist, daß Kontr 1 wie die beiden Patientengruppen zuerst die Paarvergleichsliste, dann den Wissenstest bearbeitete, daß bei Kontr 2 aber zuerst der Wissenstest als Lerninstrument über die Krankheiten, dann erst die Paarvergleichsliste vorgegeben wurde.

Die Skalenwerte für beide Gruppen nach den verschiedenen Skalierungsverfahren sowie die entsprechenden Kenngrößen finden sich in Tab. 18 und 19, die graphische Darstellung der Skala nach dem Iterationsverfahren in den Abb. 1 bzw. 6.

Die Übereinstimmungskoeffizienten für den Paarvergleich haben mit .474 bzw. .497 etwa die gleiche Größenordnung wie diejenigen der Patientengruppen. Die Medizinstudenten zeigten größere Einheitlichkeit der Beurteilung. Die Koeffizienten sind aber entsprechend ihrer absoluten Größe als durchaus hoch zu bezeichnen. Ebenso zeigen die Indices der Skalierbarkeit in der Größenordnung .9o bzw. .96 für Kontr 1 und .89 bzw. .96 für Kontr 2 an, daß auch für diese beiden Gruppen das lineare Modell sehr gut anwendbar ist.

Bei der Betrachtung der Skala von Kontr 1 ergibt sich, daß auch hier Grippe eindeutig als leichteste Krankheit eingestuft wird, und zwar deutlich abge-

Tabelle 18:

Skalenwerte der Gruppe Kontr 1

N = 61

Coefficient of Agreement U = o.474

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.688	o.ooo	o.ooo
Akne	o.5o9	1.133	1.4oo	o.734
Neur	1.394	o.942	2.666	1.746
Blin	o.594	o.682	1.921	o.7o6
Blut	o.598	o.895	1.688	o.823
Alko	1.531	o.993	2.92o	1.756
Zuck	1.182	o.638	2.531	1.4o6
Syph	1.8o6	o.934	3.3o3	2.o3o
Hirn	2.o75	o.884	3.65o	2.3oo
Tb	2.1o2	o.786	3.669	2.213
Herz	2.224	o.771	3.787	2.449
Depr	1.617	o.941	3.o6o	1.842
Quer	2.577	1.246	4.51o	2.929
Krbs	2.939	1.455	5.342	3.366
Schi	2.186	1.o3o	3.866	2.41o
Chi^2:	242.336		2o2.32o	156.235
S_D :				2.562
S_L :				72.636
S_T :				8o.223
Rho :				.9oo5
Rss :				.9631

Korrelationen (r):

Skala V zu Skala III: .985

Iteration zu Skala V : .996

Iteration zu Skala III: .989

Tabelle 19:

Skalenwerte der Gruppe Kontr 2

N = 71

Coefficient of Agreement U = o.497

	Skala V	Sigma	Skala III	Iteration
Grip	o.ooo	1.819	o.ooo	o.ooo
Akne	o.427	1.395	1.324	o.473
Neur	1.48o	o.675	3.192	1.612
Blin	o.776	o.772	2.444	o.674
Blut	o.994	o.9o4	2.586	o.994
Alko	1.639	1.o21	3.336	1.771
Zuck	1.352	o.588	3.o42	1.484
Syph	1.761	1.176	3.6o6	1.761
Hirn	2.127	1.o25	4.o71	2.259
Tb	1.879	o.535	3.613	2.o11
Herz	2.494	o.717	4.363	2.626
Depr	1.644	o.761	3.365	1.776
Quer	2.637	1.179	4.862	2.95o
Krbs	3.141	1.447	5.865	3.597
Schi	2.12o	o.985	4.o38	2.252
Chi^2:	289.73o		247.59o	183.925
S_D :				2.591
S_L :				75.966
S_T :				85.15o
Rho :				.8875
Rss :				.9649

Korrelationen (r):

Skala V zu Skala III: .981

Iteration zu Skala V : .996

Iteration zu Skala III: .974

setzt von den drei nächsten, Blinddarmentzündung, Akne und Bluthochdruck (Tab. 18). Diese nehmen als weniger schwere Krankheiten in etwa wieder eine identische Skalenposition ein. Dann folgt mit deutlicher Skalendifferenz eine Mittelgruppe von Krankheitsbildern, beginnend mit Zuckerkrankheit, mit den drei psychiatrischen Bildern Neurose, Alkoholismus, endogene Depression, welche praktisch als gleich schlimm beurteilt werden, und mit Syphilis, Tuberkulose, Hirnarteriosklerose (Verkalkung), Schizophrenie und Herzinfarkt. Wiederum eindeutig abgesetzt sind die zwei schwersten Krankheitsbilder Querschnittslähmung und Krebs. Die psychiatrischen Bilder werden also eindeutig den mittelschweren Erkrankungen zugeordnet, und sie werden sozusagen in zwei Untergruppen geteilt, Neurose, Alkoholismus, endogene Depression einerseits und Hirnarteriosklerose und Schizophrenie andererseits, welche beide in Bezug auf die subjektiv empfundene Schwere voneinander praktisch nicht unterschieden werden.

Kontr 1 gehört zu den Beurteilergruppen, die insgesamt eine deutlicher gespreizte Skala erstellen, also besser zwischen den Krankheiten differenzieren. Die Gruppierung der Krankheitsbilder zeigt sich entsprechend in Abb. 6, in der durch Elimination der verschieden großen Spreizung zwischen den Gruppen noch einmal zu sehen ist, daß Grippe eher als Bagatellkrankheit eingeordnet wird, daß dann die Gruppe der leichten Krankheiten folgt, Zuckerkrankheit aber zu den schwereren zählt, und daß als sehr schwere Krankheiten Querschnittslähmung und Krebs eingestuft werden.

Die Beurteilergruppe Kontr 2 zeigt im Vergleich zu Kontr 1 eine noch mehr gespreizte Skala, unterscheidet also noch deutlicher zwischen den Krankheitsbildern. Zugleich zeigt sich eine wenn auch nicht völlig andere Skalierung: Grippe ist wiederum die

leichteste Krankheit, Krebs die schwerste, Querschnittslähmung gehört zu den sehr schweren Erkrankungen, aber jetzt auch eher Herzinfarkt. Zu den schweren Erkrankungen (als Mittelgruppe) zählen wie bei Kontr 1 Hirnarteriosklerose (Verkalkung) und Schizophrenie mit gleichem Skalenwert, dann Tuberkulose, endogene Depression, Alkoholismus, Syphilis mit fast gleichem Skalenwert. Dicht dabei Neurose und auch Zuckerkrankheit. Wie wiederum besser in Abb. 6 zu sehen ist, wird die Mittelgruppe der Krankheiten als relativ leichter beurteilt als durch Kontr 1.

Weiter zeigt der Vergleich zwischen den Skalenwerten beider Kontrollgruppen, daß die Reihenfolge - nicht die Skalenwerte - der organischen Krankheiten im wesentlichen gleich ist. Nur bei Akne und Blinddarmentzündung skaliert Kontr 2 einen deutlicheren Abstand als Kontr 1, wodurch eine Vertauschung der numerischen Skalenpositionen zustande kommt. Auch bei den psychiatrischen Krankheitsbildern gibt es eine Umordnung, die aber nicht sehr erheblich erscheint. Hirnarteriosklerose und Schizophrenie tauschen numerisch gegenüber Kontr 1 die Position, faktisch halten sie aber eine identische Skalenposition. Deutlicher verändert gegen Kontr 1 ist die relative Position von endogener Depression und Alkoholismus zu Syphilis. Bei Kontr 2 werden diese drei Krankheiten jetzt als gleich schwer beurteilt.

4.7 Informationsstand über die Krankheiten und Kriterien der Einstufung in Relation zur Skalierung bei den Kontrollgruppen

Durch die veränderte Vorgabe der beiden Untersuchungsteile Paarvergleichsliste und Wissenstest scheinen keine wesentlichen Änderungen in der Abfolge der Anordnung der Krankheiten bewirkt zu sein, wohl aber in den Skalenwerten selber und in der Differenzie-

rung der einzelnen Krankheiten gegeneinander.

Um beide Kontrollgruppen im Wissenstest miteinander zu vergleichen, stehen die Daten in Tab. 15 und Tab. 16 zur Verfügung.

In Tab. 15 zeigt sich, daß die Gruppe Kontr 1 in den Kennwerten für die Lösungshäufigkeiten der jeweils richtigen Items anscheinend besser liegt als Kontr 2. Ein Wert liegt bei 89.6 % (Bluthochdruck), nur ein weiterer wesentlich unter 9o %, nämlich der für Grippe mit 85.7 %. Es wurde hier bereits darauf hingewiesen, daß die für Grippe ausgewählten Items sich als sehr schwierig herausstellten. Zu einer Prüfung auf Unterschiede im mittleren Leistungsniveau zwischen beiden Gruppen Kontr 1 und Kontr 2 bieten sich weiter die Daten der Tab. 16 an, welche die Verteilung der Gesamtrohwerte im Wissenstest festhalten. Der generelle Vergleich über alle mit dem Wissenstest geprüften Gruppen hinweg mit dem H-Test von KRUSKAL & WALLIS erbrachte einen sehr signifikanten Prüfwert von 11.4; die nachfolgende post-hoc Prüfung mit dem erweiterten multiplen Vergleich nach NEMENYI konnte jedoch keine Gruppenunterschiede zwischen den einzelnen Gruppen identifizieren. Wir gingen davon aus, daß der globale statistische Vergleich nicht nur auf die Lokationsunterschiede, sondern ebenso, und wohl hier wesentlich, auf die Dispersionsunterschiede, nämlich die sehr abweichende Verteilung der Gruppe Tb ansprach.

Unter Verletzung der sehr strengen Regeln von LIENERT - wir halten diese Verletzung bei unserem Problem, zumal bei dem nicht streng experimentellen Untersuchungsansatz und als weiteren post-hoc Vergleich für vertretbar - haben wir die Daten für den GRW der Gruppen Kontr 1 und Kontr 2 mit dem U-Test nach MANN & WHITNEY (DIEHL & KOHR 1977, LIENERT 1973) korrigiert für Ranggleichheiten zwischen den Gruppen nach

WALTER (LIENERT 1973) verglichen. Der hochsignifikante Wert $z_{U\text{-}corr} = 5.7$ ($p < 0.001$) zeigt einen überzufälligen Unterschied zwischen den Gruppen an. Kontr 1 erzielt deutlich häufiger bessere Gesamtwerte im Wissenstest als Kontr 2. Die relativen Häufigkeiten in der jeweils zweiten Spalte der Tab. 16 zeigen, daß 22 % der Gruppe Kontr 1, nur 14 % der Gruppe Kontr 2 Testwerte von 25 - 27 erreichen, weiter 36 % von Kontr 1, aber wiederum nur 28 % von Kontr 2 Werte zwischen 28 und 30. Anders haben 23 % von Kontr 1 niedrigere Werte zwischen 22 und 24 gegenüber 32 % von Kontr 2. Diese Unterschiede verdeutlichen die bessere Leistung in der Gruppe Kontr 1 als Ganzes.

Dieser Unterschied läßt sich nicht aus genuinen Unterschieden der beiden Gruppen erklären. In Bezug auf Alter, Schul- und Berufsausbildung unterscheiden sie sich nicht. Es zeigt sich hier vielmehr, daß nicht nur der Wissenstest bei Kontr 2 einen Lerneffekt für die Bearbeitung der Paarvergleichsliste bringt, sondern, so muß man folgern, umgekehrt das Nachdenken über die einzelnen Krankheitsbilder, welche ja sukzessiv und dann in jeder möglichen Kombination gegeneinander verglichen werden müßten, ebenso einen Lerneffekt für die Bearbeitung des Wissenstests. Eine alternative Erklärung, gegenseitige Information der Beurteiler, ist nicht völlig auszuschließen, wurde aber bei der Durchführung kaum beobachtet und jeweils zu unterbinden versucht und stünde auch im Widerspruch zu den differierenden Ergebnissen beider Gruppen.

Zu überlegen wären hier Folgerungen für eine mehrmalige Durchführung einer Paarvergleichsliste mit den gleichen Vergleichsobjekten bei identischer Beurteilergruppe. Zu erwarten wäre möglicherweise, daß sich die Skalierung zumindest bei den ersten Durchgängen verändert und erst dann stabilisiert.

Abschließend sind nun noch die Beurteilungskriterien zu nennen, welche von den beiden Gruppen im Nachgespräch genannt wurden. Die Befragung war hier nicht sehr ergiebig, da der interessierte Teil der Beurteiler (jeweils etwa zwei Drittel der Gruppe, welche man am besten wiederum in zwei Untergruppen, in passiv Interessierte und wenige aktiv Interessierte, die energisch und viel fragten, unterteilen kann) in beschränkter Zeit sehr viele Fragen zum Gesamtzusammenhang der Untersuchung hatte. Manche Teilnehmer wollten recht genau wissen, wie ihre Daten nun ausgewertet würden, wobei sich einige sehr mißtrauisch zeigten und ihre Bogen bis zuletzt behielten, um sicher zu sein, daß man ihren Fragebogen auch wirklich nicht als ihren persönlichen identifizieren könne (daher die Kategorien "keine Angabe" in Tab. 8, vor allem bei Kontr 2). Dann wurden eben auch allgemeinere Fragen zur Psychiatrie und auch zur Psychologie gestellt.

Die Kriterien, welche spontan genannt wurden, waren wieder die gleichen wie bei den übrigen Gruppen: Insgesamt ein Vorverständnis von Krankheit als vor allem somatisches Geschehen, wobei man an Heilbarkeit/Behandelbarkeit, Lebensgefahr, körperliche Beeinträchtigung als Folge denkt. Bei diesen beiden Gruppen ist - entsprechend dem Auswahlprinzip -in der Regel keine prägende Erfahrung mit einer eigenen Erkrankung anzunehmen.

4.8 Prüfung auf geschlechtsspezifische Unterschiede

Bei der bisherigen Analyse der von den einzelnen Beurteilergruppen erstellten Skalen wurde versucht, immer zugleich mögliche Einflußfaktoren mitzubeachten. Dabei kamen wesentlich vor allem Grad der Informiertheit über Krankheiten, evtl. gruppenspezifische

Kriterien der Einordnung, Erfahrung mit einer konkreten Krankheit in den Blick. Bisher nicht geprüft wurde der mögliche Einfluß des Geschlechts der Beurteiler. Es war leider nicht möglich, eine Gruppe von weiblichen Beurteilern in die Untersuchung einzubeziehen, die etwa in Bezug auf Alter und Bildungsstand mit den beiden Kontrollgruppen vergleichbar gewesen wäre.

Um den nach den wenigen bisher vorliegenden Arbeiten zum Thema der Beurteilung von Krankheiten zu erwartenden bzw. möglichen Einfluß der Geschlechtszugehörigkeit auf die Skalierung von Krankheiten wenigstens überblickshaft zu kontrollieren, wurden die Gruppen, bei denen es aufgrund der zahlenmäßigen Größe und einer einigermaßen gleichgewichtigen Verteilung der Geschlechter sinnvoll erschien, Med 1 und Alk, in Untergruppen weiblicher und männlicher Beurteiler aufgeteilt. Die Skalen, getrennt nach Geschlecht, finden sich in Tab. 2o. Sie sind mit dem Iterationsverfahren erstellt. Die Gruppengrößen sind dadurch gering. Daher ist die Beurteilung von Unterschieden der Skalenwerte einzelner Krankheiten wie auch der Vergleich zwischen den Skalen verschiedener Gruppen entsprechend vorsichtig vorzunehmen. Zugleich sind die Übereinstimmungskoeffizienten U und die Indices der Skalierbarkeit Rho und Rss aufgeführt.

Zum Vergleich der Skalen sind einmal jeweils in der Spalte "Rang" in Tab. 2o die Rangplätze aufgeführt, in der Spalte "Diff(erenz)" jeweils die Rangplatzunterschiede zwischen den Geschlechtern jeder Gruppe. Die Rangplatzdifferenzen deuten noch am meisten auf Unterschiede hin. In der Gruppe der Medizinstudenten sind sie als minimal anzusehen; allein bei der Skalierung von endogener Depression ergibt sich ein Differenzwert von 2, und zwar dadurch, daß die weiblichen Beurteiler endogene Depression höher einstuften. Dieser Effekt wird allerdings etwas bedeutsamer,

Tabelle 2o:
Einstufung der Krankheiten durch nach Geschlecht unterschiedene Untergruppen: Medizinstudentinnen (Med 1 W) und -studenten (Med 1 M) sowie weibliche (Alk W) und männliche (Alk M) Alkoholpatienten (Iteration)

	Med 1 W				Med 1 M	Alk W				Alk M
N =	25				41	17				24
	Rang	Diff	Rang			Rang	Diff	Rang		
Grip	o.ooo	1	o	1	o.ooo	o.ooo	1	o	1	o.ooo
Akne	o.21o	2	o	2	o.398	o.691	3	o	3	o.437
Neur	1.473	6	o	6	1.775	1.62o	6	o	6	1.3oo
Blin	o.41o	3	o	3	o.746	o.482	2	o	2	o.4o2
Blut	1.o67	4	o	4	1.616	o.8o3	4	o	4	o.832
Alko	1.876	9	1	1o	2.218	1.7o6	8	o	8	1.656
Zuck	1.39o	5	o	5	1.69o	1.492	5	o	5	1.235
Syph	1.553	7	o	7	2.o96	1.729	9	2	7	1.627
Hirn	2.3o6	11	o	11	2.572	2.73o	14	2	12	2.196
Tb	1.637	8	1	9	2.147	1.694	7	3	1o	1.819
Herz	2.365	12	1	13	3.o42	2.239	11	2	13	2.448
Depr	2.oo8	1o	2	8	2.1o2	2.o27	1o	1	9	1.662
Quer	2.811	14	o	14	3.847	2.828	15	1	14	2.589
Krbs	2.9o2	15	o	15	3.937	2.716	13	2	15	2.98o
Schi	2.536	13	1	12	2.982	2.69o	12	1	11	1.832
U=	.516				.582	.52o				.473
Rho=	.711				.78o	.645				.758
Rss=	.913				.959	.887				.926

Interkorrelationsmatrix der Skalen:

		Produkt-Moment-Korr. r			
		Med 1 W	Med 1 M	Alk W	Alk M
Rang-korr. R_S	Med 1 W	--	.98o4	.9737	.9662
	Med 1 M	.9857	--	.9383	.971o
	Alk W	.9571	.9357	--	.944o
	Alk M	.975o	.9821	.95oo	--

wenn man gleichzeitig wiederum die unterschiedliche Spreizung der Skala beachtet. Die männlichen Beurteiler differenzieren stärker zwischen den Krankheiten, besonders deutlich bei Med 1 M gegenüber Med 1 W, weniger bei Alk M gegen Alk W.

Wenn man nach den Gründen für die geringere Spreizung der Skalen bei den weiblichen Beurteilern sucht, sieht man in den Indices der Skalierbarkeit zwar nicht erhebliche, aber doch deutliche Unterschiede. Vor allem der Index Rho, der nach unseren Überlegungen stärker die Abweichung von der Linearität, also vom eindimensionalen Modell widerspiegelt, ist jeweils deutlich niedriger bei den weiblichen Beurteilern; .711 bei Med 1 W gegen .78o bei Med 1 M, .645 bei Alk W gegen .758 bei Alk M.

Es ist jetzt darauf zurückzukommen, daß beispielhaft bei Kontr 2 ein Zusammenhang zwischen Spreizung der Skala und Anzahl von intransitiven Urteilen nachgewiesen wurde. Auch bei den beiden nach Geschlecht getrennten Gruppen Med 1 und Alk haben wir den Unterschied in der Anzahl der zirkulären Triaden geprüft. Es zeigen sich deutliche numerische Unterschiede. Für Med 1 M beträgt die mittlere Anzahl zirkulärer Triaden ($\overline{ZT}$; arithmetisches Mittel) 6.76 bei einer Standardabweichung von 5.8, für Med 1 W ist $\overline{ZT}$ = 9.o8, σ = 6.5. Entsprechend sind die Werte für Alk M: $\overline{ZT}$ = 6.42, σ = 3.76, bei Alk W dagegen $\overline{ZT}$ = 1o.35, σ = 6.96. Unter dem Modell einer statistischen Prüfung zwischen Zufallsstichproben aus gleichen (d. h. mit gleichem Mittelwert; H_o) oder verschiedenen Grundgesamtheiten (H_1), welches man insofern heranziehen kann, als es sich zwar nicht um echte Zufallsstichproben handelt, die Gruppen aber dennoch in eingeschränkter Weise als repräsentativ für wenn auch eng umschriebene Grundgesamtheiten angesehen werden können (z. B. Medizinstudenten eines bestimmten Semesters; Alkoholkranke bestimmten Typs), wurde eine

Signifikanzprüfung mit dem U-Test von MANN & WHITNEY (mit Korrektur nach WALTER) durchgeführt. Der parametrische t-Test wurde nicht gewählt, da die Annahme einer Normalverteilung der ZT-Werte in einer zugrundeliegenden theoretischen Population, vor allem aber auch die vorauszusetzende Varianzhomogenität nicht berechtigt erscheint. Der Vergleich mit dem U-Test zeigt, daß die Unterschiede in der zentralen Tendenz bei beiden Gruppen hochsignifikant sind. Für Med 1 M gegen Med 1 W beträgt $z_{U\text{-}corr} = 8.158$ (p o.oo1), für Alk M gegen Alk W $z_{U\text{-}corr} = 4.773$ (p o.oo1).

Es gibt keinen Anlaß, da die Zahl der ZT insgesamt in Relation zur maximal möglichen Zahl sehr niedrig bleibt, für die weiblichen Beurteilergruppen, erst recht nicht bei den Medizinstudentinnen, geringere Kompetenz anzunehmen. Auch Unterschiede etwa der Altersverteilung oder im Bildungsniveau schlagen nicht zu Buche; bei Med 1 ist das Bildungsniveau "per definitionem" gleich, die Prüfung der Altersverteilung erbringt nicht-signifikante Chi^2-Werte. Im Bildungsniveau liegt Alk W gegenüber Alk M (Chi^2; Schule 1o.56, Beruf 12.o8) signifikant (o.o5) höher. Aber es gilt, daß zwischen Bildungsniveau und Anzahl intransitiver Urteile kein Zusammenhang besteht bzw. wenn, dann eher ein gegenläufiger (Nullkorrelationen oder negative Korrelationen).

Man muß also davon ausgehen, daß weibliche Beurteiler, jedenfalls in den Gruppen, die in dieser Untersuchung befragt wurden, in der Tendenz anders urteilen als männliche, indem sie nicht so streng an einem eindimensionalen Beurteilungskriterium festhalten, allerdings wiederum nicht so sehr, daß sich erhebliche Differenzen zu den männlichen Beurteilers ergeben oder das lineare Skalierungsmodell nicht mehr anwendbar wird.

Für die beiden Gruppen der Alkoholpatienten zeigen

sich bei den Rangplatzdifferenzen häufigere und größere Unterschiede, wobei aber zu beachten ist, daß eine Rangplatzdifferenz immer mindestens eine zweite bedingt.

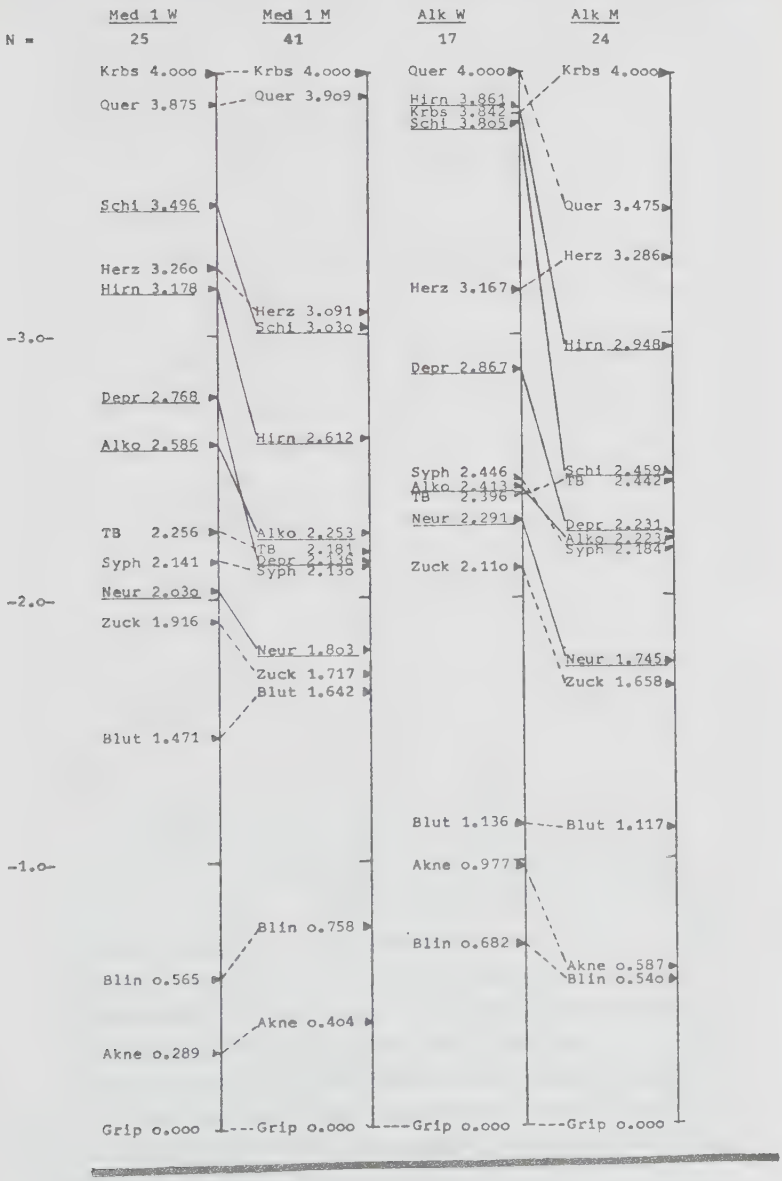
Zum weiteren Vergleich sollen die linear transformierten graphisch dargestellten Skalen (lineare Transformation wie bei Abb. 6) in Abb. 7 herangezogen werden, welche aufgrund der Elimination der verschieden starken Spreizung der Skalen einen Vergleich gegeneinander übersichtlicher machen.

Für die beiden Gruppen aus Med 1 ist der Unterschied schnell zu erkennen. Die Studentinnen bewerten die meisten psychiatrischen Erkrankungen schwerer als die männlichen Beurteiler aus der Studentengruppe, deutlich vor allem bei endogener Depression, Hirnarteriosklerose und Schizophrenie, während sie sonst mit den männlichen Beurteilern weitestgehend übereinstimmen.

Weniger einfach strukturiert sind die Unterschiede zwischen Alk W und Alk M. Aber wiederum übereinstimmend zum Vergleich zwischen den beiden Medizinergruppen ergibt sich, daß die psychiatrischen Erkrankungen von den weiblichen Beurteilern schwerer bewertet werden (jeweils in Relation zu der Krankheit mit dem höchsten Skalenwert), interessanterweise ausgenommen Alkoholismus. Zugleich ist aber auch die Skala der weiblichen Alkoholkranken quasi "gedrittelt" in eine Gruppe leichter Krankheiten, eine Gruppe eher schwerer Erkrankungen mit Zuckerkrankheit, Neurose, Tuberkulose, Alkoholismus und Syphilis und zwei "Ausreissern", endogene Depression und Herzinfarkt als wiederum noch schwerere, und dann eine dritte Gruppe, welche mit praktisch gleichem Skalenwert als die schwersten Erkrankungen das Maximum der Skala bilden, Schizophrenie, Krebs, Hirnarteriosklerose ("Verkalkung") und Querschnittslähmung.

Alk M verteilt dagegen gleichmäßiger über die gesamte

Abbildung 7:
Skalen der Gruppen Med 1 und Alk, getrennt nach Geschlecht



(transformierte Skalen; vgl. zu Abb. 6)

Skala, und auch der Abstand von Krebs als der schwersten Erkrankung zu Querschnittslähmung als der subjektiv als zweitschwerste beurteilten ist deutlich größer als bei Alk W.

Für einen Vergleich der Beurteilung durch die verschiedenen Geschlechter, soweit das aus dem vorliegenden Material zu eruieren ist, ist also als Ergebnis festzuhalten, daß es offensichtlich geschlechtsspezifische Bewertungsunterschiede gibt, daß nämlich weibliche Beurteiler psychische (psychiatrische) Erkrankungen als schwerer, schlimmer empfinden.

Es kann jedoch selbstverständlich nichts darüber gesagt werden, ob ein solcher Unterschied zwischen den Geschlechtern auch bei anderen Gruppen wiedergefunden würde. Ebenso sind Vermutungen über die Gründe für eine andere Bewertung dieser Krankheitsgruppe aufgrund des zur Verfügung stehenden empirischen Materials notwendigerweise spekulativ. Es wäre möglich, an unterschiedliche Erlebnisweisen oder Erfahrungen zu denken. Möglich wäre als Erklärung aber auch, daß eine größere Sensibilität für die quasi experimentelle Situation (Zusammenhang mit der Psychiatrie) und eine Anpassung an die vielleicht vermuteten Erwartungen des Untersuchers (Alkoholpatienten) oder der Institution (Medizinstudenten im Psychiatrie-Kurs) bestand. Was die Kriterien der Beurteilung angeht, bei Med 1 detailliert, bei Alk global erhoben, zeigen sich keine bedeutsamen Unterschiede.

Bei einer Analyse und Interpretation der Unterschiede ist schließlich für beide untersuchten Gruppen eine Relativierung angezeigt, die eine Überbewertung bestehender Unterschiede verhindert. Betrachtet man die Interkorrelationsmatrix der Skalen (Tab. 2o), und zwar sowohl der Produkt-Moment-Korrelationen r wie auch die Rangkorrelationen R_S, so sieht man, daß

die Übereinstimmung zwischen den Gruppen sowohl in Bezug auf die reine Einordnung der Krankheiten der subjektiv empfundenen Schwere nach als auch in der Gewichtung durch die Skalenwerte sehr erheblich ist. Alle Korrelationen liegen über .93, erreichen teilweise fast die Größe 1.oo. Unterschiede zwischen den Gruppen sind also in Relation zum Grad der Übereinstimmung deutlich weniger gewichtig.

4.9 Der Vergleich aller Gruppen gegeneinander: Die differierende Bewertung körperlicher und psychiatrischer Krankheitsbilder

In Tab. 21 sind die Skalenwerte der Krankheiten, berechnet nach dem Iterationsverfahren, wie sie sich aus der Beurteilung durch die sechs verschiedenen Gruppen ergeben haben, in einer Gesamtübersicht zusammengestellt. Es sind die gleichen Skalenwerte, welche auch in Abb. 1 in graphischer Form dargestellt sind.

Als erstes Ergebnis für einen Gesamtvergleich ist zu wiederholen, was sich bereits beim Vergleich einzelner Gruppen miteinander herausstellte. Die einzelnen Gruppen erstellen unterschiedlich stark gespreizte Skalen. Dabei konnte gezeigt werden, daß man diese Unterschiede dahingehend psychologisch interpretieren muß, daß die Gruppen unterschiedlich stark zwischen den Krankheitsbildern differenzieren. Es ist daran zu erinnern, daß ein deutlicher Zusammenhang besteht zwischen diesem Ausmaß der Unterscheidung zwischen den Krankheiten und einmal der Häufigkeit von Extremverteilungen wie zum anderen geringerer Zahl von intransitiven Urteilen im Paarvergleich (vgl. Abb. 4; Extremgruppenvergleich aus Kontr 2). Am stärksten differenziert Med 1, gefolgt von Kontr 2 und Kontr 1; dann folgen Med 2 und Alk, und am wenigsten unterscheidet Tb.

Tabelle 21:
Skalenwerte (Iteration) geordnet entsprechend Kontr 1

	Med 1	Med 2	Tb	Alk	Kontr 1	Kontr 2
N =	66	37	48	41	61	71
Grip	0.000	0.000	0.174	0.000	0.000	0.000
Blin	0.708	0.395	0.000	0.454	0.706	0.674
Akne	0.355	0.051	0.103	0.660	0.734	0.473
Blut	1.454	0.773	0.827	0.893	0.823	0.994
Zuck	1.660	1.185	1.298	1.435	1.406	1.484
Neur	1.784	1.369	1.113	1.486	1.746	1.612
Alko	2.230	1.938	1.594	1.781	1.756	1.771
Depr	2.186	2.171	1.270	1.992	1.842	1.776
Syph	1.957	1.544	1.444	1.806	2.030	1.761
Tb	2.129	1.541	1.670	1.841	2.213	2.011
Hirn	2.578	1.926	1.838	2.537	2.300	2.259
Schi	2.882	2.556	1.763	2.266	2.410	2.252
Herz	2.855	2.402	2.176	2.497	2.449	2.626
Quer	3.473	3.183	2.484	2.852	2.929	2.950
Krbs	3.610	3.329	2.883	3.032	3.366	3.597

Die Skalenwerte der einzelnen Krankheiten differieren zwischen den Gruppen durchaus (Abb. 1). Alle Gruppen aber beurteilen einheitlich Krebs als die schwerste, schlimmste Krankheit, Querschnittslähmung als die zweitschwerste Erkrankung der 15 vorgegebenen Krankheiten. Ebenso wird - ausgenommen von Tb - Grippe als die leichteste, am wenigsten schlimme Erkrankung empfunden, und es folgen gewöhnlich Akne, Blinddarmentzündung und Bluthochdruck, wobei drei Gruppen die Reihenfolge von Akne und Blinddarmentzündung vertauschen.

Im mittleren Bereich der Skalen, also bei den Krankheiten, welche als recht schlimm bzw. recht schwer angesehen werden, unterscheiden sich die Beurteilergruppen vor allem in der Reihenfolge der Anordnung. Weniger bedeutsam sind dagegen Differenzen der Skalenwerte gleicher Krankheiten. Die Unterschiede beziehen sich insgesamt fast ausschließlich auf differierende Einordnung der psychischen Erkrankungen.

Das wird in Abb. 8 wie auch in Abb. 9 (welche gegenüber Abb. 8 transformiert ist wie Abb. 6 gegenüber Abb. 1) deutlich erkennbar. In diesen Skalen sind die psychiatrischen Erkrankungen Neurose, Alkoholismus, endogene Depression, Hirnarteriosklerose und Schizophrenie herausgelöscht. Wie beim Vergleich zwischen jeweils zwei einzelnen Gruppen ergibt sich auch für den Vergleich aller Beurteilergruppen, daß (bis auf eine geringfügige Vertauschung bei fast gleichem Skalenwert zwischen Syphilis und Tuberkulose bei Med 2, und bis auf die oben bereits erwähnten unterschiedlichen Anordnungen bei den als wenig schlimm beurteilten Krankheiten) alle Gruppen die körperlichen Erkrankungen in einer identischen Rangfolge anordnen.

In Abb. 1o sind zum weiteren Vergleich nur die psychiatrischen Krankheitsbilder enthalten (mit Anfangs- und Endpunkt jeder Skala; nicht transformiert). Sie werden zwischen den Gruppen sehr viel uneinheit-

Abbildung 8:
Skalenpositionen der somatischen Krankheiten ohne die psychiatrischen Erkrankungen

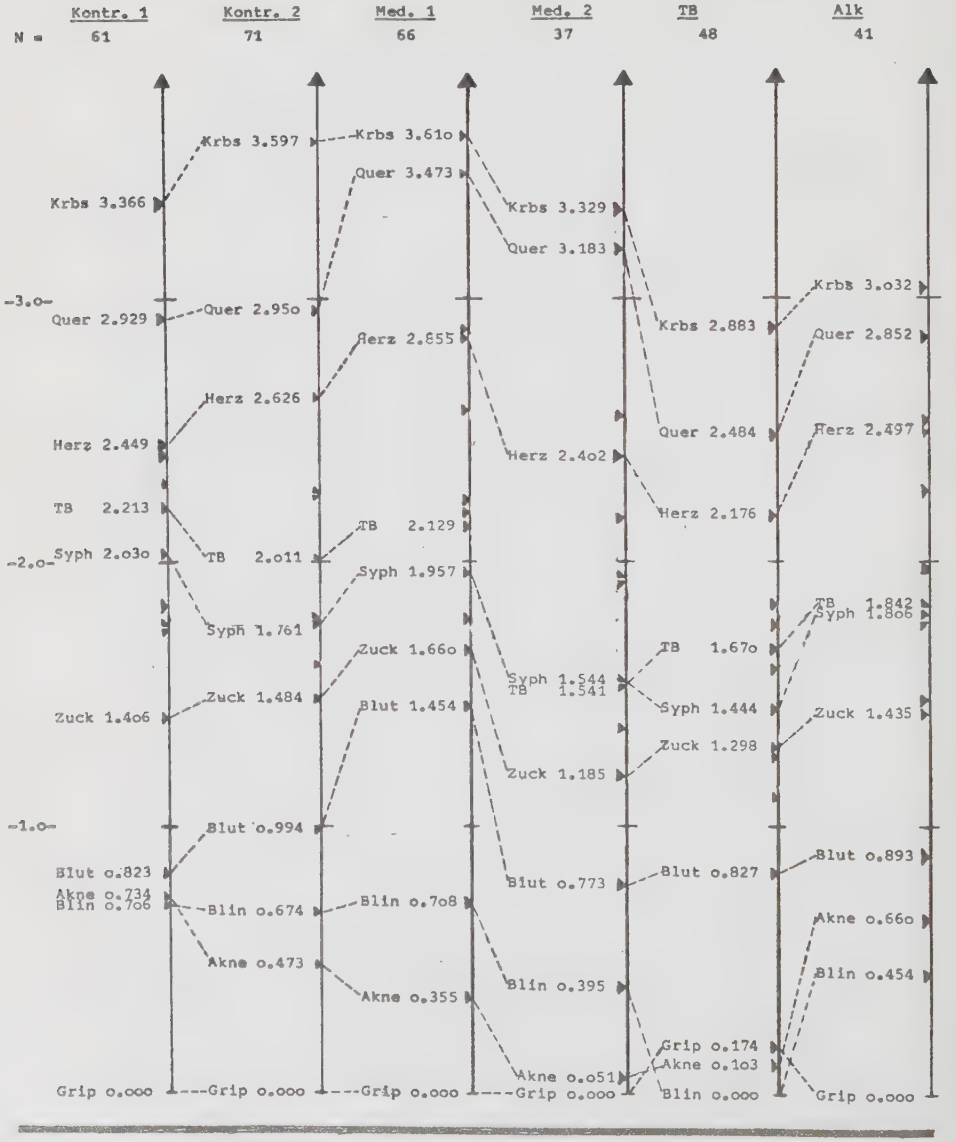


Abbildung 9:

Skalenpositionen der somatischen Krankheiten ohne die psychiatrischen Erkrankungen (transformierte Skalen)

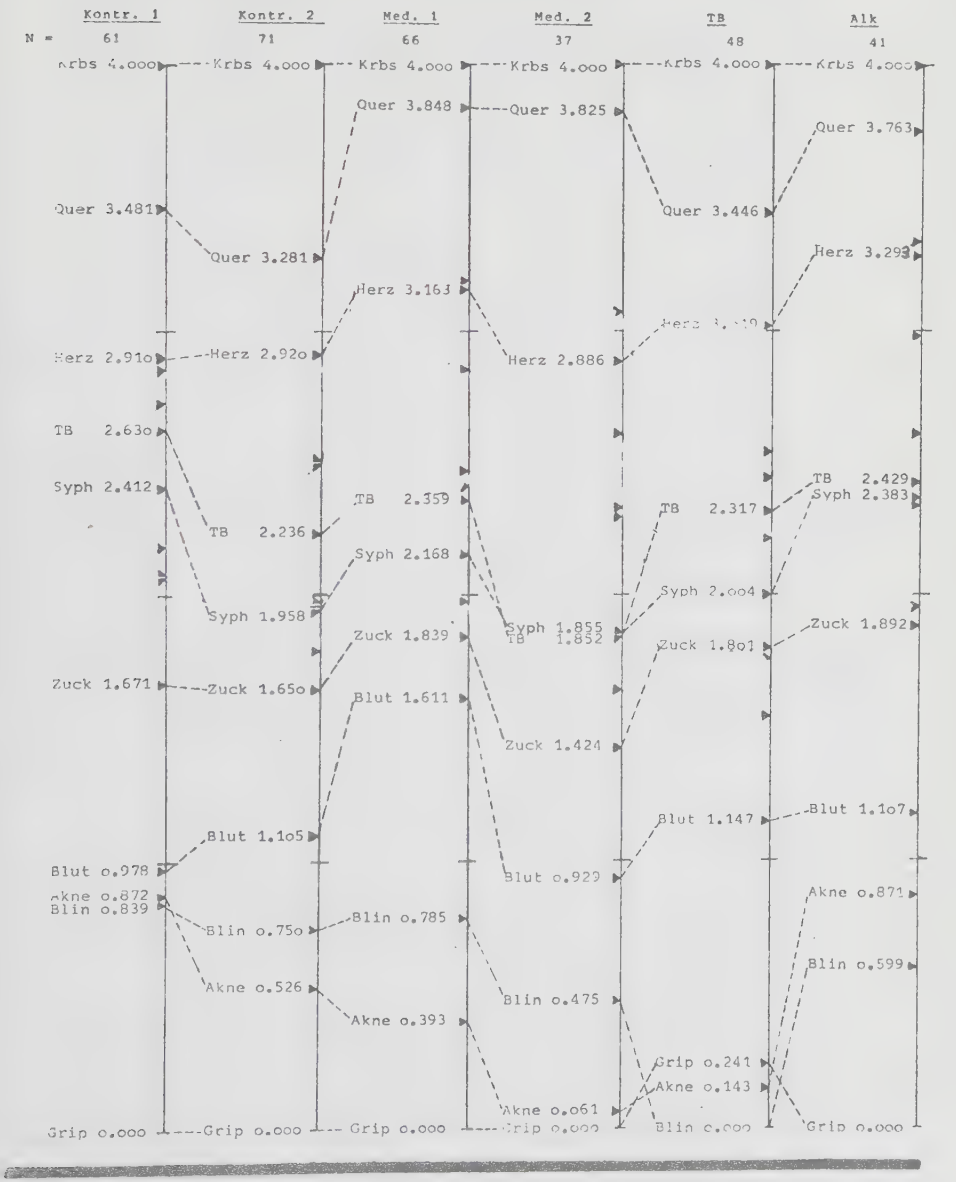
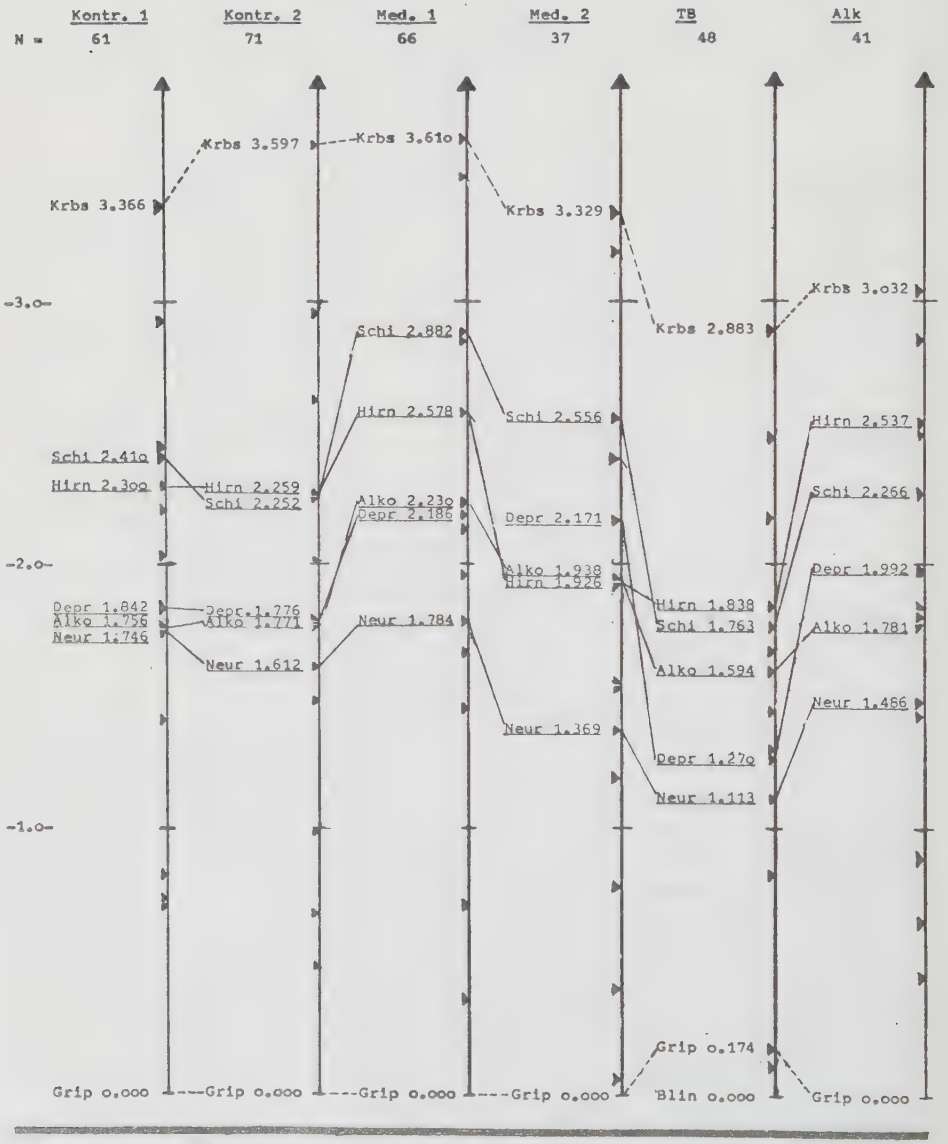


Abbildung 10:
Skalenpositionen der psychiatrischen Erkrankungen ohne die somatischen Krankheiten (ausgenommen leichteste und schwerste)



licher eingeordnet, auch wenn man sie nur in Relation zueinander betrachtet. Und im Vergleich wiederum mit Abb. 1 wird ersichtlich, daß sie auch in Relation zu den körperlichen Erkrankungen mittleren Schweregrades von den verschiedenen Gruppen sehr unterschiedlich eingeordnet werden. In der hier gewählten Anordnung der Skalen ist das an den sich in diesen Skalenbereichen häufig überkreuzenden Verbindungslinien festzustellen.

Als Maßzahlen für den Grad der Übereinstimmung der von den verschiedenen Gruppen erstellten Skalen können - wie beim Vergleich der nach Geschlecht unterteilten Gruppen Med 1 und Alk - die Interkorrelationen dienen. Bei einer Beurteilung dieser Kenngrößen ist allerdings zu bedenken, daß einmal die Produkt-Moment-Korrelation r gegen auch einseitige Linearverschiebung nur einer Skala wie auch gegen einseitige Transformation invariant ist, also die unterschiedliche Spreizung von zwei Skalen nicht widerspiegelt. Sie spricht nur auf die unterschiedlichen relativen Distanzen der Objekte, damit eingeschlossen auf unterschiedliche relative Anordnung an. Die ebenfalls bestimmten Rangkorrelationen nach SPEARMAN (R_S) sprechen nur auf unterschiedliche Anordnung der Krankheiten, also auf Veränderungen der Rangfolge an, bleiben aber gegenüber Linearverschiebungen, Transformationen und auch Änderungen der Distanzen (Differenzen) invariant. Schließlich zeigen die Korrelationen der Skalen der parallelen Gruppen aus Med 1 und Med 2 (die 29 Studenten, welche an beiden Durchgängen teilnahmen; Abb. 2) mit ihrer jeweiligen Gesamtgruppe, daß bei einer Heranziehung als Kenngröße für den Skalenvergleich im Zusammenhang dieser Untersuchung die Höhe der Korrelation anders zu beurteilen ist als im üblichen Zusammenhang. Obwohl beide Korrelationen (r) fast 1.oo erreichen (.995 und .997), sind dennoch Unterschiede festzustellen.

In Tab. 22 finden sich die Interkorrelationen der vollständigen Skalen mit allen 15 Krankheiten aus Tab. 21, oberhalb der Diagonalen die Produkt-Moment-Korrelationen r, unterhalb die Rangkorrelationen R_S.

Die Produkt-Moment-Korrelationen bewegen sich zwischen etwa .95 und .99 mit $\bar{r}$ = .972. Sie zeigen eine weitgehende Übereinstimmung zwischen den Skalen der verschiedenen Beurteilergruppen an, allerdings unter den oben aufgeführten Einschränkungen, daß sie die unterschiedliche Spreizung der Skalen eliminieren. Trotz aller Kautelen ist also die Übereinstimmung sehr hoch.

Auch die Rangkorrelationen spiegeln diese Übereinstimmung wider, zeigen in ihrer größeren Schwankung zwischen etwa .89 und .99 aber eher die Unterschiede der Anordnungen, wobei sie allerdings minimale numerische Skalendifferenzen überbewerten, die ja als Unterschiede in die Rangordnung genauso eingehen wie sehr große Differenzen.

Die Unterschiede zwischen den verschiedenen Skalen der Beurteilergruppen werden erst deutlich, wenn man die Skalenwerte getrennt für die körperlichen und die psychiatrischen Krankheitsbilder interkorreliert. In Tab. 23.1 und 23.2 sind die Korrelationen einmal auf der Basis der 1o nicht-psychiatrischen Krankheiten oberhalb der Diagonale, dann auf der Basis der 5 psychiatrischen unterhalb der Diagonale gegenübergestellt (Produkt-Moment-Korrelationen r).

Der Vergleich der oberen Dreiecksmatrix aus Tab. 23.1 (1o Items) mit der oberen Dreiecksmatrix in Tab. 22 (alle 15 Items; r) zeigt, daß sich die Koeffizienten in der gleichen Größenordnung bewegen, daß aber die "1o-er"-Korrelationen insgesamt geringfügig höher sind, in jedem einzelnen Fall zumindest gleich hoch. Die Durchschnittswerte betragen bei 15 Items $\bar{r}$ = .972 (s.o.), bei 1o Items $\bar{r}$ = .982.

Tabelle 22:
Interkorrelationsmatrix der Skalen (Iteration) der sechs Beurteilergruppen (oberhalb der Diagonalen: Produkt-Moment-Korrelationen r; unterhalb: SPEARMAN-Rang-Korrelationen R_S)

R_S — r	Med 1	Med 2	Tb	Alk	Kontr 1	Kontr 2
Med 1	—	.983	.969	.978	.969	.981
Med 2	.982	—	.951	.962	.953	.965
Tb	.939	.889	—	.963	.956	.977
Alk	.961	.936	.939	—	.98o	.978
Kontr 1	.961	.936	.961	.979	—	.987
Kontr 2	.985	.943	.957	.986	.982	—

$\bar{r}_{15} = .972$

Tabelle 23:
Interkorrelationsmatrizen der auf 1o (organische Krankheiten) bzw. auf 5 (psychiatrische Krankheiten) Skalenwerte reduzierten Skalen der sechs Beurteilergruppen (oberhalb der Diagonalen jeweils: Basis 1o Skalenwerte; unterhalb: Basis 5 Skalenwerte)

Tabelle 23.1
Produkt-Moment-Korrelationen r

r_5 r_{1o}	Med 1	Med 2	Tb	Alk	Kontr 1	Kontr 2
Med 1	--	.989	.971	.985	.971	.988
Med 2	.854	--	.978	.977	.968	.986
Tb	.887	.592	--	.979	.964	.98o
Alk	.859	.631	.842	--	.99o	.99o
Kontr 1	.915	.644	.8o1	.875	--	.988
Kontr 2	.939	.653	.898	.948	.976	--

$\bar{r}_5 = .859$ / $\bar{r}_{1o} = .982$

Tabelle 23.2
Rangkorrelationen R_S

R_{S5} R_{S1o}	Med 1	Med 2	Tb	Alk	Kontr 1	Kontr 2
Med 1	--	.983	.952	.988	.988	1.ooo
Med 2	.6oo	--	.939	.976	.976	.983
Tb	1.ooo	.6oo	--	.964	.964	.952
Alk	.8oo	.4oo	.8oo	--	1.ooo	.988
Kontr 1	.9oo	.7oo	.9oo	.9oo	--	.988
Kontr 2	.8oo	.4oo	.8oo	1.ooo	.9oo	--

Die Produkt-Moment-Korrelationen auf der Basis von 5 psychiatrischen Items in der unteren Dreiecksmatrix von Tab. 23.1 sind hingegen gegenüber den Korrelationen auf der Basis der 1o organischen Erkrankungen deutlich niedriger; ihr Mittelwert beträgt $\bar{r}$ = .859. Diese Unterschiede zeigen übereinstimmend mit der Analyse der graphischen Skalen, daß die verschiedenen Beurteilergruppen sich weitgehend einig sind - mit geringen Ausnahmen - in ihrer subjektiven Einschätzung darüber, wie schlimm die körperlichen Erkrankungen sind. Bei den psychiatrischen Erkrankungen hingegen besteht eine deutliche Urteilsunsicherheit. Es gibt gemeinsame Einschätzungen: Neurose wird von allen als die leichteste der psychischen Erkrankungen empfunden, aber nur bei Med 1 und Med 2 sowie Kontr 1 erhält Schizophrenie den Rang der schwersten psychischen Erkrankung, dabei mit deutlicher Skalendifferenz zu den anderen psychischen Krankheiten nur bei den Medizinern. Bei den anderen Gruppen ist es Hirnarteriosklerose. Im Blick auf die Beurteilungskriterien kann man sagen: Mit hoher Wahrscheinlichkeit deswegen, weil der "organische" Akzent schon im Begriff enthalten ist. Auch die übrigen Krankheiten "tauschen" sozusagen von Gruppe zu Gruppe die Reihenfolge und in geringerem Maße auch die Skalenpositionen.

In gleicher Weise zeigen die Rangkorrelationen R_S diese Unterschiede, allerdings ausschließlich die in der Anordnung der Krankheiten auf den Skalen. Für alle 15 Krankheiten bewegen sie sich etwa in der gleichen Größenordnung (Tab. 22, untere Dreiecksmatrix) wie die Produkt-Moment-Korrelationen. Für die zehn nicht-psychiatrischen Erkrankungen (Tab. 23.2, obere Dreiecksmatrix) erreichen die Rangkorrelationen ebenfalls die gleiche Höhe wie die entsprechenden Produkt-Moment-Korrelationen (Tab. 23.1, obere Dreiecksmatrix). Geringfügig deutlicher noch machen sie

die Unterschiede in der Einordnung der psychiatrischen Erkrankungen (Tab. 23.2, untere Dreiecksmatrix).

Wie sind nun Unterschiede der einander entsprechenden Beurteilergruppen Med 1 und Med 2 sowie Kontr 1 und Kontr 2 zu beurteilen? Die Korrelation der Gesamtskalen von Med 1 und Med 2 beträgt .983, für die reduzierte Skala nur mit den 1o körperlichen Erkrankungen (r_{1o}) .989, für die fünf psychiatrischen alleine nur .854. Das zeigt, daß die Vermehrung der Information durch den Psychiatriekurs, welche Med 1 und Med 2 unterscheidet, deutliche Veränderungen in der relativ persönlichen, subjektiven Einschätzung der Krankheiten bewirkt. Es resultiert daraus, daß vor allem Hirnarteriosklerose relativ weniger schlimm bewertet wird, dann auch Neurose; endogene Depression wird hingegen deutlich schlimmer, Schizophrenie und Alkoholismus wiederum werden nur geringfügig niedriger eingestuft.

Wie hat sich der Versuch ausgewirkt, bei Kontr 2 gegenüber Kontr 1 durch den zu Anfang vorgegebenen "Wissenstest" sachliche Informationen (wenn auch relativ knapp) als wirksamen Faktor einzuführen? Für die Skala aller 15 Krankheiten zeigt die Korrelation mit .985 eine hohe Übereinstimmung, bei den 1o organischen Krankheiten alleine keine nennenswert größere (.988). Aber auch bei der Einschätzung der psychiatrischen Erkrankungen sind bei diesen beiden Beurteilergruppen keine großen Differenzen zu erkennen, wie die Korrelation (r_5) mit .976 anzeigt. Es kann also nicht davon ausgegangen werden, daß diese Information vor der Beurteilung im Paarvergleich eine wesentlich andere Bewertung als bei den nicht zusätzlich informierten Gruppen bewirkt hat, auch nicht in Bezug auf die psychiatrischen Erkrankungen. Der Effekt des Lerntests geht also vor allem dahin, daß die Krankheiten insgesamt deutlicher unterschieden werden, was in der zweitgrößten Spreizung der Skala resultiert.

4.1o Psychologische Charakteristik typischer Beurteilergruppen und Beurteiler: Kontrastgruppenvergleich und kontrastierend urteilende Individuen

4.1o.1 Kontrastierende Gruppen

Bisher wurde versucht, Bedingungen des Urteilsverhaltens, also der Skalierung, aufgrund von Randdaten zu analysieren. Das sind etwa Gruppenzugehörigkeit und - damit teilweise identisch - Grad der Informiertheit über Krankheiten (Mediziner vs. andere bzw. Punktwert im Wissenstest), oder Geschlechtsunterschiede. Es ist zu fragen, ob man noch eingehendere Aufschlüsse über Bedingungen des Beurteilerverhaltens erreichen kann, wenn man in der Beurteilung von Krankheiten deutlich sich unterscheidende Untergruppen oder auch einzelne Beurteiler miteinander vergleicht. Bei einem solchen Vergleich kann man abheben einmal auf mitgegebene Randdaten (Geschlecht, Ergebnis im Wissenstest, Bildungsstand u. ä.), oder aber auf immanente Daten, etwa bestimmte Gruppierungen von Krankheitsbildern auf der Skala der subjektiven Schwere.

Die Frage läßt sich mit der Bildung von sogenannten "Extremgruppen", anders benannt "Kontrastgruppen", beantworten. Im Zusammenhang mit der Skalierung durch Paarvergleichsurteile ist diese Methodik von SÜLLWOLD (1969 b) allerdings in einem anderen thematischen Kontext angewandt worden. Es ist zuerst zu beschreiben, wie die Gruppen kontrastierender Beurteiler definiert wurden, da es eine generell festgelegte Regel nur insoweit gibt, als man üblicherweise Gruppen vom unteren oder oberen Ende einer Verteilung gegeneinander stellt. Man wählt z. B. das Drittel mit den niedrigsten Werten in einer zu definierenden Kenngröße und stellt es dem Drittel mit den höchsten Werten in dieser Kenngröße gegenüber, um die Gruppen dann im Hinblick auf andere Merkmalsausprägungen zu vergleichen.

Aus allen Gruppen gemeinsam - ausgenommen die Medizinstudenten - wurden die kontrastierenden Gruppen wie folgt gebildet: Nach den individuellen Dominanzmatrizen wurden die Rangplätze der einzelnen Krankheiten für jeden Beurteiler errechnet, indem bestimmt wurde, wie oft jeder einzelne Beurteiler ein Krankheitsbild als "schlimmer" ansah als die 14 anderen. Dann wurden die Beurteiler entsprechend den aufsteigenden Rangplätzen für jeweils ein bestimmtes Krankheitsbild geordnet. So lassen sich zwei kontrastierende Gruppen bilden, welche sich darin unterscheiden, daß die Mitglieder der einen Gruppe einem Krankheitsbild implizit die höchsten Rangplätze zuordnen, während die kontrastierende Gruppe dem gleichen Krankheitsbild nur die niedrigsten Rangplätze zuteilt.

Kontrastgruppen in Bezug auf die Einstufung aller 15 Krankheitsbilder zu bilden und dann für jede Einzelgruppe die Skala der Krankheiten zu bestimmen, hätte allerdings einen sehr großen Rechenaufwand mit sich gebracht (15 x 2 = 3o Skalierungen). Wenn dann weiter aus jeder der sechs ursprünglichen Untersuchungsgruppen jeweils zwei kontrastierende Gruppen gebildet worden wären, wäre der Aufwand noch einmal auf das Sechsfache gestiegen. Vor allem aber ist der Wert solcher Kontrastgruppen - Skalierung sehr gering einzuschätzen, weil die Unsicherheit der Skalenwerte indirekt proportional ist zur Gruppengröße, aufgrund derer die Skala bestimmt wird. In dieser Überlegung ist begründet, daß bei der Bildung von Kontrastgruppen alle Gruppen mit Ausnahme der beiden Medizinergruppen (Med 1 und Med 2) zusammengefaßt wurden. Gerechtfertigt ist es, weil man diese vier Gruppen als im wesentlichen durch das Alter unterschieden sehen muß, weil sie sich aber als Gruppen in der Einschätzung der Krankheitsbilder, wie die bisherigen Vergleiche ergaben, nicht fundamental unterscheiden.

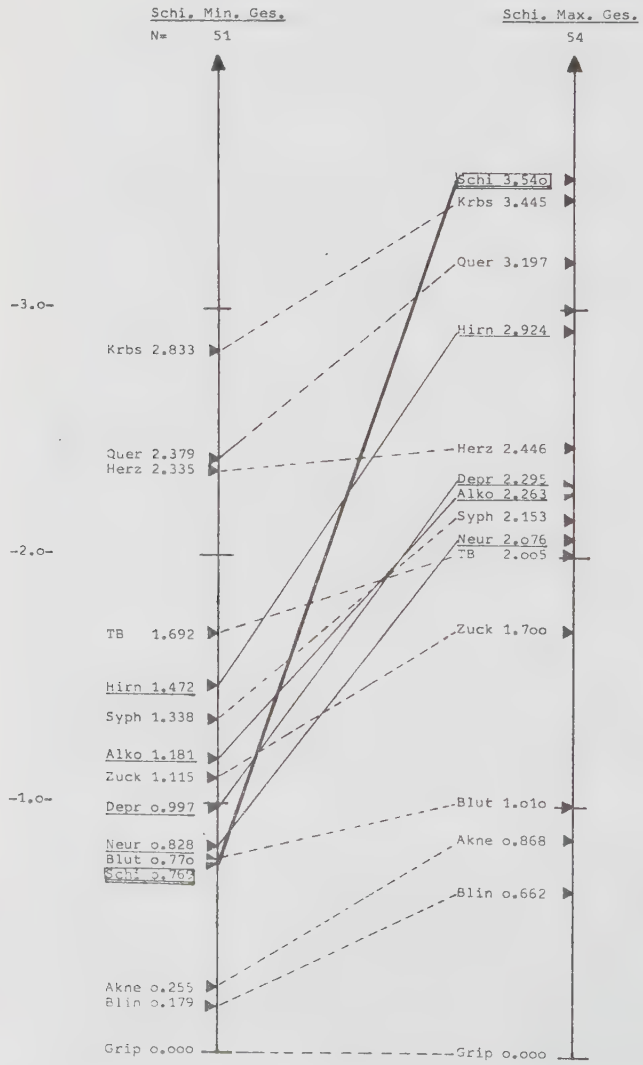
Aus dieser Gesamtgruppe aus Kontr 1, Kontr 2, Tb und Alk wurden Extremgruppen nur in Bezug auf die fünf besonders interessierenden psychischen Erkrankungen gebildet. Konkret geschah das so, daß im ersten Schritt jeweils die unteren bzw. oberen 25 % (um den Kontrast zu erhöhen) mit der niedrigsten bzw. höchsten Bewertung jeweils eines Krankheitsbildes bestimmt wurden (von insgesamt 221 Beurteilern die "untersten" und die "obersten" 55). Dann wurden aus diesen kontrastierenden Gruppen alle diejenigen ausgeschieden, welche dem in Frage stehenden Krankheitsbild einen Rangplatz implizit zugeordnet hatten, der auch noch in der jeweils übrig bleibenden Restgruppe der mittleren 5o % vertreten war. Wenn also etwa bei der Einstufung des Krankheitsbildes "Schizophrenie" der noch zum ersten Quartil gehörende 55. Beurteiler für diese Krankheit den Rang 8 zugeteilt hatte, der nächstfolgende 56. aber auch noch den gleichen Rangplatz 8, so wurden aus den unteren 25 % aller Beurteiler zusätzlich diejenigen, welche Schizophrenie mit dem Rang 8 eingeordnet hatten (also achtmal im Paarvergleich "bevorzugt" hatten) ausgeschieden. Dadurch ist erklärt, daß in der Gruppe "Schi.Min." 51 Beurteiler mit den implizit erzielten Rangplätzen 1 bis 7 übrig blieben (o, d. h. keine Bevorzugung, kam bei Schi. nicht vor). Entsprechend wurde bei der Bildung der kontrastierenden Gruppe "Schi.Max." verfahren, sowie bei der Auswahl der übrigen Kontrastgruppen.

Die Abb. 11 zeigt die Skalen der beiden so gebildeten Extremgruppen. Von der Gruppe "Schi.Min.Ges." mit 51 Beurteilern wird Schizophrenie fast am unteren Ende der Skala angeordnet, von der kontrastierenden Gruppe, welche Schizophrenie fast immer als die schwerere von beiden jeweils im Paar zu vergleichenden Krankheitsbildern bezeichnete, wird im Gruppenurteil die gleiche Krankheit an die Spitze der Skala gesetzt.

Abbildung 11:

Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Schizophrenie

(alle Beurteiler ohne vorherige Prüfung der Kenntnis des Krankheitsbildes (Ges.))



Kennwerte:		
U = =	.44o	.557
Rho =	.884	.8o8
Rss =	.959	.953

Es war jedoch noch ein weiterer Auswahlschritt notwendig. Bei einer ersten Betrachtung dieser Gruppe zeigte sich nämlich, daß in der Gruppe Schi.Min.Ges. eine große Anzahl von Beurteilern vertreten war, welche im Wissenstest die Items zu Schizophrenie nicht richtig beantwortet hatte. Man muß daraus folgern, daß Unkenntnis über diese Krankheit mit ein bedeutender Faktor für eine in Bezug auf subjektiv empfundene Schwere geringe Einstufung war. Zugleich darf man generalisieren, daß dies wohl bei der Beurteilung aller Krankheiten gilt. Da aber nicht Kontraste analysiert werden sollten, welche auf Unkenntnis und vielleicht daraus resultierenden zufallsmäßigen bzw. "ratenden" Einstufungen beruhten, wurden in einem dritten Schritt bei allen psychischen Krankheitsbildern, deren Einstufung als Grundlage für die Bildung von Kontrastgruppen benutzt worden war, diejenigen Beurteiler ausgeschieden, welche die zugehörigen Items des Wissenstests zum jeweiligen Krankheitsbild nicht richtig beantwortet hatten.

Die Zusammensetzung dieser endgültig resultierenden Kontrastgruppen ist in Tab. 24 dargestellt. Es findet sich dort jeweils eine Einzeltabelle für die Gruppen, welche ein psychiatrisches Krankheitsbild gering ("Min.") oder hoch ("Max.") eingestuft haben. In den ersten Zellen der beiden korrespondierenden Tabellenhälften ist aufgeführt, über welche Rangplätze die Einstufungen in diesen Kontrastgruppen streuen, so z. B. für Neur.Min. von Rang 1 bis 4 (Rang o, d. h. keine Bevorzugung kam nur bei Beurteilern vor, welche im letzten Auswahlschritt als weniger kompetent ausgeschieden wurden), bei Neur.Max. entsprechend am oberen Ende der Rangplatzverteilung, nämlich von Rang 9 bis 12 (13 und 14 als mögliche Ränge wurden aber auch hier nicht vergeben).

Zu beachten ist, daß einmal bei den Gruppen Hirn.Min. und Schi.Min. die Streubreite der Ränge im Vergleich zu

Tabelle 24:
Kennwerte der Kontrastgruppen

		NEUR			ALKO			HIRN			DEPR			SCHI		
	Ränge:	1 - 4			o - 4			1 - 7			2 - 5			1 - 7		
		Ges	-	N	Ges	-	N	Ges	-	N	Ges	-	N	Ges	-	N
MIN.	Kontr 1	11	1	1o	11	4	7	16	3	13	16	2	14	12	3	9
	Kontr 2	14	5	9	16	3	13	22	6	16	16	5	11	14	7	7
	Tb	15	5	1o	6	4	2	8	4	4	16	9	7	14	8	6
	Alk	1o	3	7	8	2	6	7	2	5	6	o	6	11	7	4
		5o	14	36	41	13	28	53	15	38	54	16	38	51	25	26
	$\bar{x}_{GRW}$	22.62		25.83	21.23		23.93	22.66		25.o5	22.o4		24.45	19.84		23.85
	s_{GRW}	6.13		3.1o	6.6o		4.55	5.9o		3.59	6.8o		4.93	6.89		4.72
	$\bar{x}_{ZT}$	1o.6o		8.94	11.37		9.82	14.55		13.oo	1o.41		7.21	16.39		13.oo
	s_{ZT}	1o.46		11.32	7.59		7.64	12.35		12.56	9.51		6.73	13.75		12.34

		9 - 12			1o - 14			13 - 14			1o - 14			12 - 14		
		Ges	-	N	Ges	-	N	Ges	-	N	Ges	-	N	Ges	-	N
MAX.	Kontr 1	13	1	12	12	o	12	9	4	5	11	2	9	14	o	14
	Kontr 2	11	1	1o	16	4	12	12	2	1o	11	2	9	16	2	14
	Tb	6	1	5	17	5	12	3	o	3	6	4	2	11	o	11
	Alk	4	1	3	1o	o	1o	9	2	7	13	2	11	13	o	13
		34	4	3o	55	9	46	33	8	25	41	1o	31	54	2	52
	$\bar{x}_{GRW}$	25.15		26.5o	24.o5		26.o2	25.o6		25.96	23.88		25.45	26.19		26.54
	s_{GRW}	4.95		2.89	6.39		3.99	4.21		3.96	5.1o		3.38	3.9o		3.24
	$\bar{x}_{ZT}$	12.26		12.5o	12.36		1o.o7	7.36		5.88	1o.1o		1o.97	7.94		7.62
	s_{ZT}	11.19		11.41	13.21		9.19	6.39		4.63	8.o7		8.81	6.45		6.17

den anderen sehr groß ist (Rang 1 bis 7), was mit Sicherheit die bereits erwähnte Urteilsunsicherheit bei diesen Krankheiten auch in diesen Gruppen widerspiegelt. Zum zweiten ist anzumerken, daß bei Neur.Max. bei den N = 3o endgültigen Beurteilern die Rangplätze 13 und 14 auch nicht vergeben wurden. Insgesamt ist zu sehen, daß die Gruppen deutlich kontrastieren. In den folgenden Zeilen der Tabellen finden sich dann die Angaben, wieviel Beurteiler aus jeder Ursprungsgruppe zum jeweiligen unteren oder oberen Quartil der Verteilung nach Rangplätzen gehören (Kontr 1, Kontr 2, Tb und Alk unter der Spaltenüberschrift Ges. = Gesamt), wieviel aus jeder Ursprungsgruppe weiterhin ausgeschieden wurden, weil sie die zugehörigen Items des Wissenstests nicht hinreichend beantworten konnten (unter der folgenden mit "-" bezeichneten Spalte) und wieviel Beurteiler aus jeder Ursprungsgruppe schließlich übrig blieben (unter "N").

Bei Betrachtung der jeweiligen Summen in der fünften Zeile der Untertabellen fällt sofort auf, daß in den Min.Gruppen ein im Verhältnis deutlich höherer Anteil von Beurteilern ausgeschieden werden mußte, weil man ihn als weniger kompetent anzusehen hatte. In den Max.-Gruppen sind es nur wenige, zumal bei Neurose und Schizophrenie, von denen man infolge dieser Überlegungen annehmen muß, daß sie ein so akzentuiertes Urteil eher auf der Grundlage einer gewissen mangelnden Kenntnis des Krankheitsbildes abgeben.

Diese "Regel": "Irrtümliche Ansichten über psychische Krankheiten führen zu einer eher geringen Einschätzung der Krankheiten auf der subjektiven Schwere-Skala" dürfte sich besonders bei der Gruppe Tb auswirken. Bereits bei der Zusammenstellung der Ursprungsgruppen mußte hier eine hohe Zahl wegen unvollständiger Bearbeitung des Paarvergleichs ausgeschieden werden. Diese Gruppe Tb hat auch im Wissenstest insgesamt schlechtere Ergebnisse. Es ist zu sehen, daß Probanden aus

der Gruppe Tb in den Extremgruppen eher überproportional vertreten sind, zumal in den Min.-Gruppen, und daß ein hoher Anteil von ihnen die einschlägigen Items des Wissenstests nicht richtig zugeordnet hatte und deswegen ausgeschieden wurde.

Schlägt sich diese Tendenz "weniger gut Informierte urteilen eher unterschätzend (vielleicht aber andererseits auch eher überschätzend)?" auch ablesbar in den eingeführten Kennwerten für Informiertheit über die Krankheiten (GRW des Wissenstests) und Schlüssigkeit des Urteilsverhaltens im Paarvergleich (ZT) nieder? Bei einem Überblick über die mit $\bar{x}_{GRW}$ (Mittelwert der Leistung im Wissenstest, unter Ges. derjenige für die kontrastierenden Gruppen einschließlich der falsch antwortenden, unter N derjenige für die Restgruppe, deren Gesamturteil dann skaliert wurde) bzw. $\bar{x}_{ZT}$ gekennzeichneten Zeilen der Tab. 24 ist festzustellen, daß bei den Min.-Gruppen diese Mittelwerte regelmäßig niedriger liegen, die zugehörige Standardabweichung (s_{GRW}) ist regelmäßig deutlich höher. Wenn man beachtet, daß jeweils bis zu 5o % der Beurteiler ausgeschieden wurden, ist der Unterschied nicht erstaunlich. Es zeigt sich aber, daß es sich nicht nur um Beurteiler handelt, die bei den spezifischen Items nicht orientiert sind.

Für die Mittelwertsunterschiede bei der Anzahl der intransitiven Urteile wird eine Tendenz nicht so deutlich. Aber auch hier deuten die Unterschiede - kleine Durchschnittswerte in den Restgruppen - darauf hin, daß die "Regel" im Sinne der bisherigen Ergebnisse zu bestätigen ist und weiterhin dahin zu ergänzen ist, daß weniger gut Informierte auch eher inkonsistent urteilen.

Bei den Gruppen, welche psychische Krankheiten eher hoch einstufen, zeigen sich analoge Differenzen kaum. Für eine Annahme, daß möglicherweise der Schweregrad

auch überschätzt wird, finden sich keine stichhaltigen Hinweise.

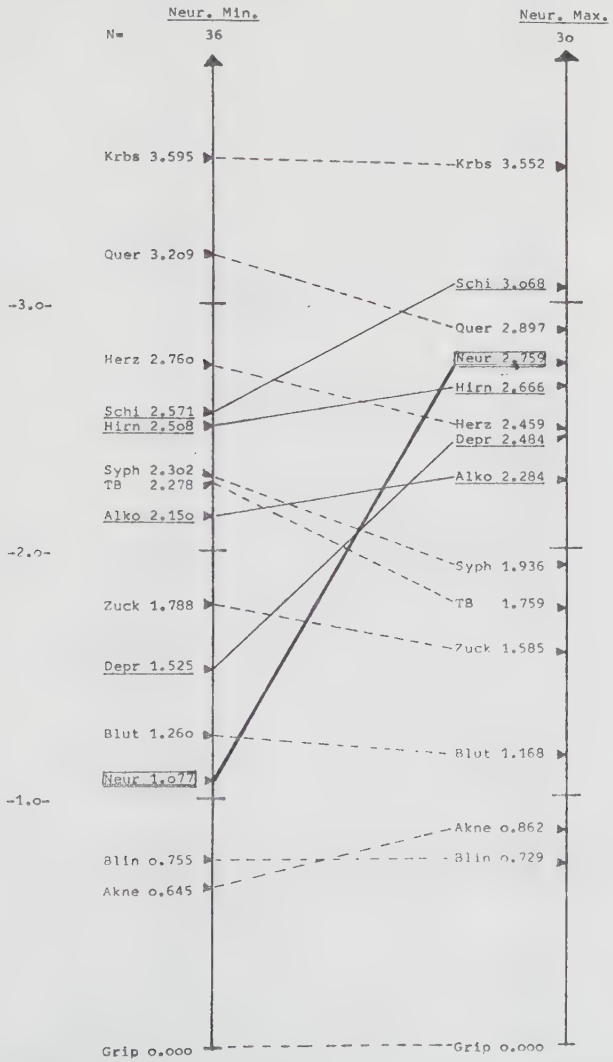
Für die Kontrastgruppen, welche Schizophrenie entweder sehr niedrig oder sehr hoch einstufen, wurden die Skalen sowohl für die ursprünglichen Gruppen (Schi.Min.Ges., Schi.Max.Ges.; Abb. 11) als auch für die endgültigen, um die weniger kompetenten Beurteiler reduzierten Kontrastgruppen Schi.Min. und Schi.Max. (Abb. 12.5) bestimmt. Die Skalen Schi.Max.Ges. und Schi.Max. sind fast identisch, da nur zwei Beurteiler ausscheiden mußten, während Schi.Min.Ges. und Schi.Min. deutlich differieren. Schizophrenie wird bei der Gesamtgruppe (Schi.Min.Ges.) als leichteste von allen psychischen Erkrankungen deutlich extrem niedrig eingestuft, fast gleich mit Bluthochdruck. Die Gruppe Schi.Min. hingegen stuft Schizophrenie nicht einmal unter den psychischen Erkrankungen als leichteste ein, sondern in einer mittleren Position. Weiterhin unterscheiden sich beide Skalen dadurch, daß Schi.Max. deutlich stärker differenziert (Spreizung der Skala). Fast identisch ist hingegen bei beiden Gruppen die Abfolge der körperlichen Erkrankungen. Nur Akne und Blinddarmentzündung haben die relativen Positionen getauscht. Diese Differenzen untermauern die Folgerungen aus dem tabellarischen Vergleich der Kontrastgruppen in ihrer vorläufigen und endgültigen Zusammensetzung. Bei extremer Einstufung eines Krankheitsbildes - diese allgemeine Aussage erscheint gerechtfertigt - ist als wesentlicher Faktor ein Mangel an Information zu vermuten, wobei solch ein Mangel trotz einer insgesamt hinreichenden Informiertheit, wie aus den bisherigen Ergebnissen zu erkennen war, am ehesten spezifisch bei psychischen Erkrankungen zu erwarten ist.

Bei der Analyse der Skalen, welche aus dem Urteil der kontrastierenden Gruppen resultieren, sollen zuerst diejenigen betrachtet werden, welche das jeweilige psychiatrische Krankheitsbild niedrig einstufen.

Es ist jetzt zu berücksichtigen, daß die weniger kompetenten Beurteiler ausgeschieden sind. In einem Überblick über die Abbildungen 12.1 bis 12.5 ist zu sehen, daß keine der Min.-Gruppen die jeweilige Krankheit niedriger als alle anderen anordnet, daß aber das herausgehobene psychiatrische Krankheitsbild, ausgenommen bei Schi.Min., bei den psychischen Erkrankungen den jeweils niedrigsten Skalenwert erhält.

In der Einschätzung der körperlichen Erkrankungen stimmen diese Gruppen im wesentlichen überein, zumal was die als schwerer eingestuften Krankheitsbilder Krebs, Querschnittslähmung und Herzinfarkt angeht. Änderungen der Reihenfolge und der Skalierung ergeben sich nur in geringerem Maße bei den als leicht eingestuften Krankheitsbildern. Als Gesamtergebnis ist festzuhalten: Für das Gruppenurteil resultiert nicht eine extrem niedrige Einschätzung der psychiatrischen Krankheitsbilder, sondern eine Einstufung im Mittelbereich. Weiterhin stimmen die fünf Gruppen mit niedrigerer Einstufung jeweils einer psychischen Krankheit untereinander im Gesamturteil weitgehend überein. Ihre Skalen stimmen aber auch mit den ursprünglichen Skalen der Gruppen Kontr 1 und 2, Tb und Alk weitgehend überein. Um das weiter zu belegen, sind in Tab. 25 die Produkt-Moment-Korrelationen der Kontrastgruppen-Skalen untereinander (Tab. 25.1) sowie mit den Ursprungsgruppen (Tab. 25.2) als Meßwerte für Übereinstimmung aufgeführt. Mit Ausnahmen, auf die noch einzugehen sein wird, ist in Teil 1 der Tabelle zu erkennen, daß die Skalierungen durch die Min.-Gruppen untereinander jeweils höher übereinstimmen ($\bar{r}$ = .937) als sich Übereinstimmungen mit den Max.-Gruppen finden lassen ($\bar{r}$ = .848). Entsprechend stimmen auch die Max.-Gruppen untereinander gut überein ($\bar{r}$ = .934), und erwartungsgemäß zeigen die jeweils korrespondierenden Kontrastgruppen, etwa Neur.Min. zu Neur.Max. die geringsten Übereinstimmungen ($\bar{r}$ = .801). Aus der Höhe der Korre-

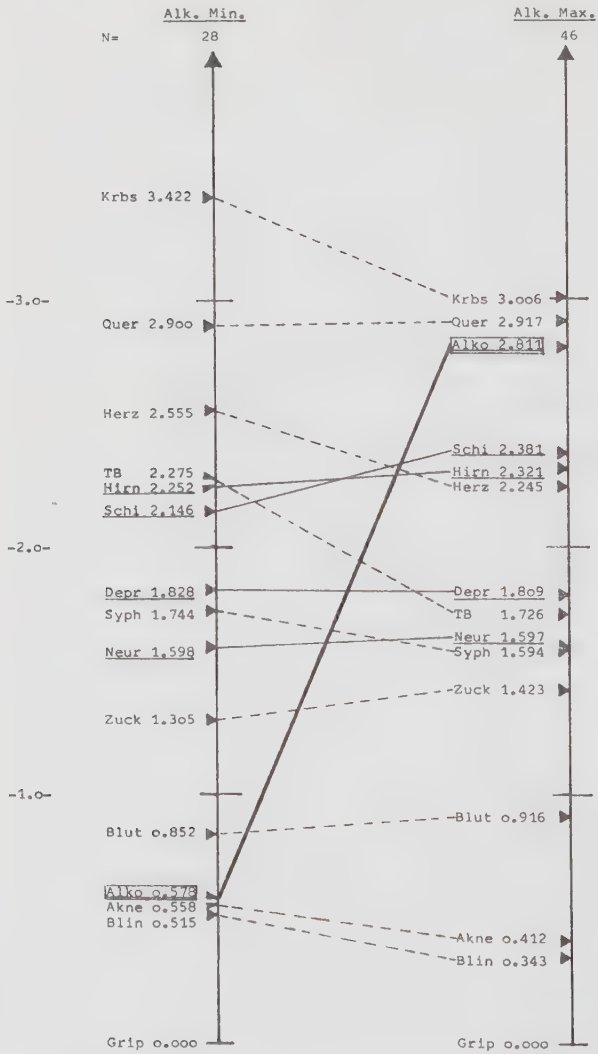
Abbildung 12.1:
Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Neurose



Kennwerte:

U =	.560	.532
Rho =	.768	.772
Rss =	.942	.937

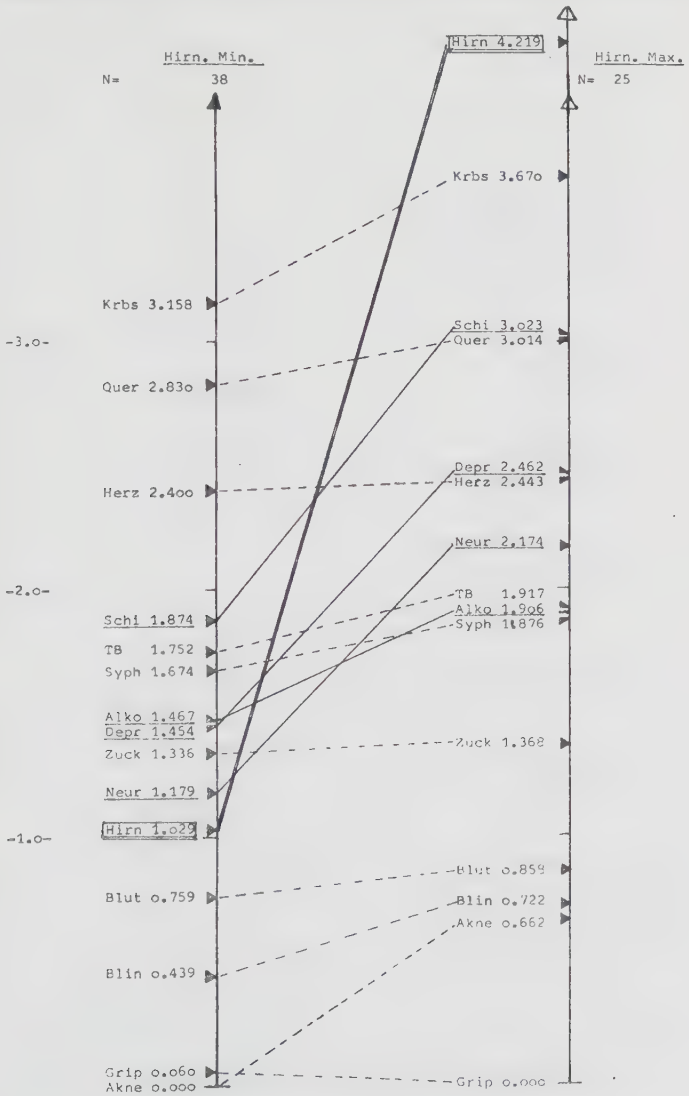
Abbildung 12.2:
Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Alkoholismus



Kennwerte:

U =	.527	.521
Rho =	.816	.842
Rss =	.942	.956

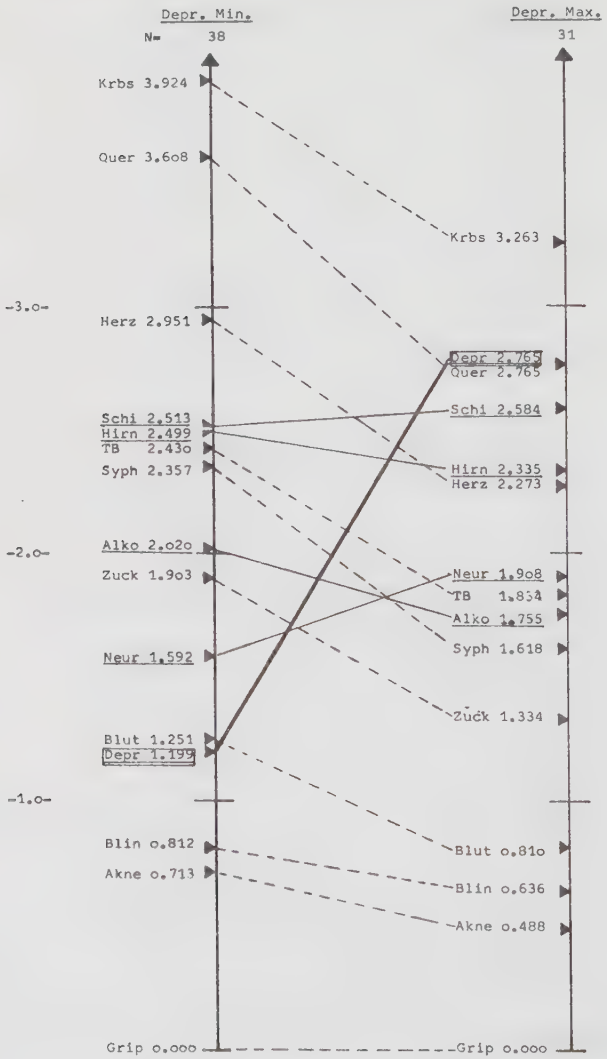
Abbildung 12.3:
Skalen der kontrastierenden Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Hirnarteriosklerose



Kennwerte:

U =	.5o9	.588
Rho =	.775	.825
Rss =	.93o	.953

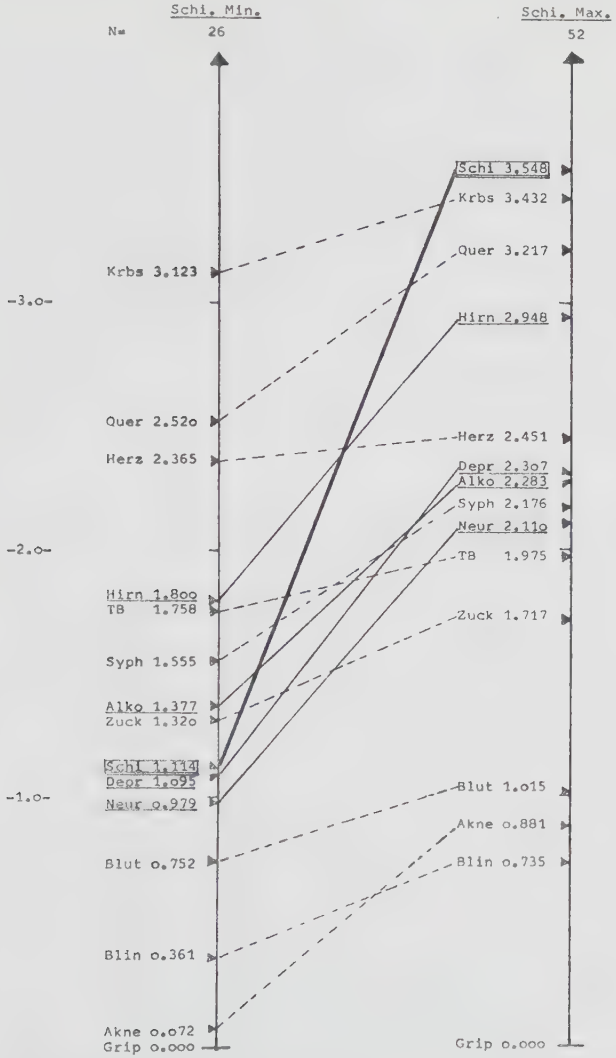
Abbildung 12.4:
Skalen kontrastierender Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von (endogene) Depression



Kennwerte:

U =	.556	.522
Rho =	.839	.751
Rss =	.962	.922

Abbildung 12.5:
Skalen kontrastierender Gruppen mit niedriger (Min.) vs. hoher (Max.) Einstufung von Schizophrenie



Kennwerte:

U =	.484	.562
Rho =	.773	.789
Rss =	.933	.949

Tabelle 25:
Interkorrelationen der Extremgruppen-Skalen

25.1: Extremgruppen untereinander:

	Alk. Min.	Hirn. Min.	Depr. Min.	Schi. Min.	Neur. Max.	Alk. Max.	Hirn. Max.	Depr. Max.	Schi. Max.
Neur.Min.	.892	.928	.982	.947	.839	.918	.836	.861	.904
Alk.Min.		.895	.912	.907	.847	.786	.856	.899	.863
Hirn.Min.			.939	.940	.828	.870	.723	.869	.840
Depr.Min				.961	.836	.892	.814	.836	.881
Schi.Min.	($\bar{r}_{min}$=.937)				.791	.863	.787	.831	.805
Neur.Max.						.921	.919	.962	.966
Alk.Max.		($\bar{r}_{Gr}$=.801)					.867	.908	.937
Hirn.Max.								.919	.937
Depr.Max.	($\bar{r}_{mm}$=.848)			($\bar{r}_{max}$=.934)					.951

25.2: Extremgruppen und ursprüngliche Beurteilergruppen:

	Kontr 1	Kontr 2	Tb	Alk
Neur.Min.	.957	.965	.964	.954
Alk.Min.	.950	.945	.907	.928
Hirn.Min.	.932	.953	.948	.901
Depr.Min.	.959	.967	.960	.935
Schi.Min.	.929	.959	.964	.919
Neur.Max.	.929	.918	.875	.931
Alk.Max.	.923	.929	.939	.945
Hirn.Max.	.899	.886	.849	.932
Depr.Max.	.945	.938	.896	.956
Schi.Max.	.950	.928	.899	.958

lationen als Maß der Übereinstimmung oder Verschiedenheit ist aber wiederum herauszulesen - wie bereits beim Vergleich der ursprünglichen Gruppen untereinander -, daß die Übereinstimmung über die Einschätzung von Krankheiten auch unter kontrastierenden Gruppen die Unterschiede überwiegt.

Von noch größerem Interesse ist Teil 2 von Tab. 25. Schon der Vergleich der Skalenwerte der Min.-Gruppen mit den Skalen der ursprünglichen Gruppen wies darauf hin, daß es wenig relevante Unterschiede gibt. Die Korrelationen nun der Min.-Gruppen-Skalen zu den Ursprungsskalen liegen jeweils deutlich höher als die der Max.-Gruppen-Skalen zu den Ursprungsskalen, etwa Neur.Min. zu Kontr 1: .957, Neur.Max. zu Kontr 1: .929. Ausnahmen finden sich nur bei Alk.Min. und Alk. Max. zu Tb und Alk und bei Schi.Min./Schi.Max. zu Kontr 1 und Alk.

Generell kann das Ergebnis so beschrieben werden: Der Vergleich der Skalen der Min.-Gruppen mit den Ursprungsskalen, besonders der Vergleich der Korrelationen als Maß der Übereinstimmung zeigt, daß es keine deutlich kontrastierenden Gruppen gibt, welche psychische Krankheiten niedrig einstufen und sich dadurch vom Urteil der Ursprungsgruppen (denen wir eine gewisse Repräsentanz für die von ihnen vertretenen Gesamtpopulationen zuerkennen müssen) wesentlich unterscheiden. Kontrastierende Gruppen sind also eigentlich nur diejenigen, welche psychische Krankheiten eher als schwer einschätzen. Weiterhin kann man aufgrund der Ergebnisse nicht von fünf auch untereinander verschiedenen Kontrastgruppen (je nach Krankheitsbild) sprechen, sondern nur von einer einzigen Gruppe, die alle psychiatrischen Erkrankungen generell als eher schwer im Verhältnis zu somatischen Erkrankungen einstuft.

Die Kontrastgruppenanalyse erbringt also drei Ergebnisse:

1. Eine Einschätzung von psychischen Erkrankungen als wenig schlimm geht bei den hier befragten Beurteilern vor allem mit einer weniger guten Informiertheit über die betreffende Krankheit, wie auch der hier zum Vergleich vorgestellten Krankheiten einher.
2. Schaltet man den Faktor "weniger gute Information" aus, d. h. nimmt man aus den Kontrastgruppen alle die heraus, welche im Wissenstest ihre hinreichende Kenntnis nicht belegen konnten, so läßt sich in Relation zu den Ursprungsgruppen aller befragten Beurteiler keine echte Kontrastgruppe mehr finden, welche subjektiv den Schweregrad seelischer Krankheiten eher gering einschätzt. Insofern kann man weitergehend schließen, daß es keine dahingehend kontrastierende Meinung bzw. Einstellung bei unseren Beurteilern gab, die dazu führte, im Vergleich zur Gesamtheit der Beurteiler seelische Erkrankungen eher als leicht anzusehen.
3. Es gibt jedoch sehr wohl eine Gruppe von Beurteilern, die seelische Krankheiten in Relation zu den meisten körperlichen konstant als schwerwiegender einschätzt und so sich von der Mehrheit der Beurteiler abhebt. Da es zudem aber keine Untergruppe gibt, welche nur ein psychisches Krankheitsbild herausstellt, sondern da alle psychischen Krankheiten nahe beieinander liegende Skalenpositionen in allen diesen Gruppen erhalten, folgt, daß gerade das Moment des Psychischen für eine solche Beurteilung entscheidend ist.

4.1o.2 Kontrastierend urteilende Einzelne

Typisch für die Beurteiler, welche das Krankheitsbild Schizophrenie extrem niedrig einschätzen, ist ein Proband aus der Gruppe Kontr 2. Er ist Bundeswehrrekrut, hat weder eine abgeschlossene Hauptschulbildung noch eine Berufsausbildung, urteilt bei der Paarvergleichstechnik der Krankheiten insgesamt sehr inkonsistent

(61 ZT). Er erreicht mit 18 Punkten im Wissenstest in Anbetracht der geringen Schwierigkeit dieses informellen Verfahrens einen relativ geringen Wert (einen Prozentrang von 15 bei einem Median für Kontr 2 von 24 und einem Q1 von 21). Die richtigen Antwortitems zu geläufigen Krankheiten findet er, so bei Krebs, Querschnittslähmung, Herzinfarkt, Zuckerkrankheit, Blinddarmentzündung, Bluthochdruck und Akne. Aber auch bei Schizophrenie, Hirnarteriosklerose, Neurose, Alkoholismus und Depression antwortet er richtig. Eine wohl weniger geläufige Krankheit wie Tuberkulose kennt er nicht sicher genug. Bei den Items zu Grippe versagt er, weil hier, wie bereits diskutiert, die Itemformulierung zu schwierig war. Er ordnet die Krankheiten in folgender individueller Rangreihe an (Rang in Klammern): als leichteste Schizophrenie (o), dann Grippe (1), Zuckerkrankheit (2), Akne (3), dann Blinddarmentzündung (4.5) und Neurose (4.5) gleichrangig, Bluthochdruck (6.5) und endogene Depression (6.5) gleichrangig folgend, dann Tuberkulose (8) und Querschnittslähmung (9), und als schwerste, wieder auf gleichem Rangplatz Hirnarteriosklerose (12), Alkoholismus (12), Syphilis (12), Herzinfarkt (12) und Krebs (12).

Man kann bei diesem Beurteiler insgesamt auf eine große Urteilsunsicherheit schließen. Es dürften Zusammenhänge zum mangelnden Bildungsniveau und vermutlich höchstens gerade durchschnittlicher Intelligenz bestehen. Dennoch ist er nicht völlig falsch orientiert, kann im Wissenstest wohl eher vermutend einiges richtig zuordnen. Extrem beurteilt er Schizophrenie, wohl deshalb, weil er mit diesem Krankheitsbild keine rechten Vorstellungen verbindet.

Ein weiterer Beurteiler, der nach den Auswahlprinzipien ebenfalls der Kontrastgruppe Schi.Min. zugeordnet wurde, unterscheidet sich deutlich vom Obigen. Dieser ist ein 22-jähriger Rekrut aus der Gruppe Kontr

1, mit Abitur und abgeschlossenem Lehrberuf. Diese Kombination läßt auf eine Ausbildung im kaufmännischen oder technischen Bereich schließen, die er evtl. durch ein folgendes Fachstudium ergänzen wird. Er urteilt im Paarvergleich sehr konsistent (5 ZT) und im Wissenstest erreicht er alle möglichen 3o Punkte; er gehört also zu den gut Informierten. Er ordnet die Krankheiten folgendermaßen: Grippe (o) als leichteste, dann Akne (1), gleichrangig Bluthochdruck (2.5) und Neurose (2.5), weiter gleichrangig Blinddarmentzündung (5), Alkoholismus (5) und endogene Depression (5), dann Schizophrenie (7), Tuberkulose (8), Zuckerkrankheit (9), Syphilis (1o), Hirnarteriosklerose (11), Krebs (12), Herzinfarkt (13), Querschnittslähmung (14) als schwerste. Hier ist eine andere Grundlage des Urteilens als mangelnde Information offensichtlich, wobei man natürlich nicht sagen kann, das sei jetzt ein "richtiges" Urteil, da es - wie herausgestellt worden ist - einen objektiven Maßstab nicht gibt. Krankheiten, welche als psychisch zu klassifizieren sind, werden von diesem Beurteiler den leichteren zugeordnet: Neurose, Alkoholismus, Depression und Schizophrenie zusammen mit Grippe, Akne, Bluthochdruck und Blinddarmentzündung. Bei Hirnarteriosklerose dürfte er folglich eher von der "organischen" Konnotation ausgehen. Sie ist für ihn die viertschwerste Erkrankung, nach Querschnittslähmung, Herzinfarkt und Krebs. Aufschlußreich erscheint, daß er nicht wie die übergroße Mehrheit Krebs als schwerste Krankheit ansieht, sondern Querschnittslähmung. Man kann daraus auf seine Urteilskriterien schließen: Das Moment des Körperlichen schätzt er schwerwiegender ein als das des Psychischen, und beim Körperlichen scheint der Aspekt einer dauerhaften Behinderung (Querschnittslähmung) wiederum schwerwiegender als eine Krankheit, mit der man generell eine Lebensbedrohung assoziiert.

Bei der Darstellung der Gruppen, welche psychische Krankheiten zu den schlimmeren oder schlimmsten rech-

nen, wurde bereits herausgehoben, daß alle psychischen Erkrankungen, abgesehen von der Sonderstellung von Hirnarteriosklerose, gemeinsam gesehen werden. Ein typischer Vertreter dieser Gruppe ist ein männlicher Beurteiler aus der Gruppe Kontr 1, der wiederum sehr konsistent urteilt (9 ZT). Er ist 21 Jahre alt, hat das Abitur, noch keine Berufsausbildung. Im Wissenstest hat er nur bei den Grippe-Items (wie aufgrund der Itemformulierung sehr viele andere) Schwierigkeiten. Er ordnet die Krankheiten folgendermaßen: Als leichteste Blinddarmentzündung (o), gleichrangig Grippe (1.5) und Zuckerkrankheit (1.5), folgend Bluthochdruck (3), Herzinfarkt (4), Akne (5), Tuberkulose (6), Syphilis (7), Querschnittslähmung (8), Krebs (9), und alle fünf psychiatrischen Erkrankungen am Schluß als die schlimmsten: Alkoholismus (1o.5) und Hirnarteriosklerose (1o.5) gleichrangig, dann Neurose (12), endogene Depression (13) und Schizophrenie (14). Mit dieser Einordnung der psychiatrischen Erkrankungen ist er allerdings auch unter der Gruppe derer, die den psychischen Erkrankungen eher hohe Skalenpositionen zuordnen, ein Extremfall. Das entscheidende Kriterium für ihn ist sicher, daß es sich um seelische Erkrankungen handelt. Eine solche Rangordnung der Krankheiten findet sich ähnlich noch mehrfach, und zwar immer bei Beurteilern, welche Schizophrenie als die schlimmste Krankheit überhaupt ansehen, die sich außerdem im Wissenstest gut informiert zeigen und ein hohes Bildungsniveau haben.

Man kann sagen, daß sich hier eine andere typische Einstellung zu psychischen Krankheiten zeigt. Man könnte sie als die eines "Intellektuellen" charakterisieren, im Kontrast zu den oben charakterisierten anderen Einstellungstypen, dem (seltenen) des "Technikers" und dem des "schlecht Informierten".

5 Diskussion der Ergebnisse

Es gibt im deutschen Sprachbereich (vgl. JAECKEL & WIESER 197o, STUMME 1975; für den anglo-amerikanischen Bereich das Übersichtsreferat von RABKIN 1972) fast zahllose Untersuchungen über die Einstellung verschiedener Gruppen zum psychiatrisch oder psychisch Kranken, insbesondere zum "Geisteskranken" oder Schizophrenen. Es gibt nur wenige Arbeiten über das Selbstbild von psychiatrischen Patienten, genauer über deren eigenes Stereotyp vom Schizophrenen oder ähnliches (vgl. RABKIN 1972, VAN QUEKELBERGHE 1976). Das wesentliche Ergebnis dieser letzten Gruppe von Untersuchungen ist, daß Patienten genauso urteilen wie nicht Betroffene. Vor allem hieraus wird deutlich, daß alle diese Untersuchungen gemeinsam haben, daß sie Krankheit oder den Kranken - auch eben wenn selber Betroffene nach ihrem Stereotyp gefragt werden - gewissermaßen aus Distanz beurteilen lassen.

Die Hauptfrage dieser Arbeit ist: Wie beurteilen Probanden aus verschieden charakterisierten Gruppen, also unter verschiedenen Bedingungen, psychiatrische Erkrankungen nach dem Grad ihrer Schwere ("Schlimmheit")? Die Urteile wurden dabei nicht isoliert erfragt, sondern in Relation zu körperlichen Erkrankungen, und sie sollten abgegeben werden unter der in der Instruktion vorgegebenen Vorstellung, die Beurteiler selber erkrankten an der einen oder der anderen der jeweils paarweise vorgestellten Krankheiten.

Eingangs wurde dazu erläutert, daß hiermit aus einem Gesamtzusammenhang von Fragen, welcher in klinisch-psychologischer Praxis verankert ist, eine erste, bewußt eingegrenzte Fragestellung herausgegriffen wird und daß versucht wird, diese einer empirisch abgesicherten Klärung zuzuführen.

Der Gesamtkomplex der Fragen kann benannt werden als

"Einstellung zur (insbesondere psychiatrischen) eigenen Krankheit und deren Bedeutung für den Behandlungserfolg". Die Art der Einstellung eines Patienten zu seiner Krankheit ist von erheblicher Bedeutung für den Erfolg einer Behandlung. So muß er sich zuerst einmal selber als wirklich betroffen identifizieren. Das scheint zumal bei seelischen Erkrankungen, aber wohl auch bei mancher körperlichen Erkrankung, der klinischen Erfahrung nach schwer zu sein. Es wird bei manchen psychiatrischen Erkrankungen - ob bei allen mit Recht oder mit Unrecht ist eine weitere Frage - auch von vielen Behandelnden als oft unmöglich angesehen. Erinnert sei an die langwährende Diskussion unter Psychiatern, ob etwa bei schizophrenen Erkrankungen mangelnde Krankheitseinsicht nicht sogar zum Wesen der Krankheit gehöre (vgl. den Lexikonartikel von PAULEIKHOFF & MESTER). Hinzuweisen ist aber auch auf die Schwierigkeiten und Bedenken, welche Behandelnde damit haben, einem Patienten seine Diagnose mitzuteilen. Diese Problematik schlägt sich einerseits immer wieder auch in wissenschaftlichen Untersuchungen (nicht nur für das Gebiet der Psychiatrie) nieder, die zu klären versuchen, wie eine solche Mitteilung wirken kann (LINDEN 1969). Sie hat aber auch ihren Niederschlag in neuesten gerichtlichen Entscheidungen gefunden, wo gerade für das Gebiet der Psychiatrie die sonst inzwischen festgeschriebene Verpflichtung des Arztes, Einsicht in die Krankenunterlagen zu geben, nicht prinzipiell aufrecht erhalten wurde.

Diese Rahmenbedingungen der hier vorgelegten Untersuchungen sind deswegen noch einmal in aller Kürze dargestellt, weil nun daranzugehen ist, die referierten Ergebnisse kritisch zu würdigen. Das soll unter einem doppelten Gesichtspunkt geschehen: Zuerst soll diskutiert werden, wieweit im Blick auf die Methodik der Paarvergleichstechnik und auf das Untersuchungs-

design allgemeingültige Folgerungen möglich sind. Dann soll auf die mehr inhaltliche Bedeutung der Ergebnisse eingegangen werden. Dabei ist einmal zu fragen, wieweit die Ausgangsfrage einer Klärung zugeführt werden konnte, zum anderen, welche offenen Fragen sich möglicherweise anschließen.

5.1 Kritische Bewertung der Methode

5.1.1 Die Paarvergleichstechnik und die spezifische mathematisch-statistische Auswertung

Mögliche Skalierungsmodelle lassen sich nach verschiedenen Aspekten unterscheiden, und zwar

- einmal im Hinblick auf die Technik der Datenerhebung, und damit implizit im Hinblick auf die psychologischen Vorannahmen über die Beurteiler bzw. den Urteilsprozeß,
- zum zweiten im Hinblick auf die Grundlegung eines statistischen Modells für die Verarbeitung der Daten zu einer Skala,
- zum dritten, und hier eher innerhalb der einzelnen Modelle, im Hinblick auf verschiedene eher rechentechnische Bestimmungsverfahren (welche z. T. auch verschiedene psychologische Implikationen haben).

Wenn man alternativ zu dem hier verwendeten Skalierungsmodell andere betrachtet, so sieht man, daß die durch die paarvergleichsspezifische Art der Datenerhebung gewonnenen Daten auch alle anderen Skalierungsmodelle ermöglichen. Man ist nicht beschränkt auf das Modell des "Law of Comparative Judgement" nach THURSTONE (1927). Umgekehrt gilt das aber nur sehr eingeschränkt (GUILFORD 1954 führt verschiedene Umrechnungsmöglichkeiten, allerdings von eingeschränkter Brauchbarkeit für andere Datenerhebungstechniken an). Das ist darin begründet, daß beim Paarvergleich bereits das elementare Stadium der Meßoperation erfaßt wird.

Vor allem aber gestatten die ungleich beliebteren alternativen Methoden*), etwa Rangordnungs- oder Rating-Verfahren, nicht eine detaillierte empirische Überprüfung der meßtheoretischen Voraussetzungen, noch der individuellen Urteilskonsistenzen wie der linearen Skalierbarkeit und damit nicht die indirekte und direkte Prüfung auf die Angemessenheit eines linearen Modells. Das ist deswegen der Fall, weil sie insgesamt Urteilstransitivität "erzwingen", also weil etwa ein Rangordnungsverfahren keine zirkulären Triaden zuläßt. Der oft beschriebene Nachteil der Paarvergleichsmethode, die zeitraubende Datenerhebung, ließ sich durch die Art der Vorgabe in Form einer Liste abmildern. Dabei wird allerdings vorausgesetzt, daß eine für jeden einzelnen Beurteiler nach Zufall permutierte Anordnung der Vorgabe nicht unbedingt notwendig ist.

Zur Wahl des zugrundezulegenden Meßmodells, hier zu den Verteilungsannahmen des Law of Comparative Judgement, nämlich daß die diskriminativen Prozesse sich normal-verteilen, gibt es keine theoretische Vorentscheidung (GIGERENZER 198o, SIXTL 1967). Die empirische Prüfung besteht in der Bestimmung der Güte der Anpassung, d. h. der generellen Prüfung der Daten auf Verträglichkeit mit dem Modell, hier mittels der Methode des Index der Skalierbarkeit. Insgesamt konnte eine sehr gute Anpassung an das lineare Modell für alle Gruppen nachgewiesen werden, so daß auf eine zusätzliche Erprobung noch anderer Modelle (SIXTL 1967) mit guten Gründen verzichtet werden konnte.

Innerhalb der Paarvergleichsmethode bleiben jedoch einige Randprobleme ungelöst.

---

*) Die weniger häufig gewählte Paarvergleichstechnik wurde unserer Kenntnis nach zuletzt umfangreich von FEGER (1972) angewendet; bei ihm erscheint allerdings die Anwendung des Verfahrens von GULLIKSEN & TUKEY, welche ja eine varianzanalytische Analogie hat, nicht ganz schlüssig.

Als erstes dieser Probleme ist die Behandlung von an sich modellunverträglichen Extremverteilungen zu nennen. Die bessere Lösung schien zu sein, sie als nicht beobachtete Daten zu behandeln. Eine konkurrierende Lösung wäre gewesen, einen in Relation zu den anderen z-Werten maximalen z-Wert zu wählen. Das von uns adaptierte Rechenprogramm (Deutsches Rechenzentrum, Autor KOLBE) für Fall V und Fall III sieht vor, für p = 1.oo bzw. o.oo, z = +3.o bzw. -3.o zu setzen.

Eine probeweise Durchführung des Iterationsverfahrens mit dieser Setzung für die Beurteilergruppe Med 1 erbrachte für den Index Rho (der u. E. der kritischere der beiden Indices der Skalierbarkeit ist) einen deutlich besseren Wert von o.8972 gegen o.8156 bei unserer Lösung, während der Index für Rss mit o.969o sich kaum von o.96o5 (vgl. Tab. 1o) unterscheidet. Bei einer solchen Lösung schließen sich aber zwei Folgefragen sogleich an: Warum 3.o und nicht etwa 2.9 oder 3.5? Das wäre durch erneutes Wiederholen der Iteration mit diesen anderen Werten und Bestimmung der Kennwerte Rho bzw. Rss zu klären. Wie ist es aber dann, das ist die zweite Folgefrage, mit entsprechenden Werten bei den anderen Gruppen? Muß man dann der Vergleichbarkeit wegen am gleichen z-Wert für p = 1.oo festhalten, oder muß man sich dort ebenfalls durch wiederholte Versuche nach dem besten Index richten? Diese Fragen führen auf eine Grenze des Meßmodells. Wiederholte Versuche stoßen zugleich an eine Grenze eines sinnvoll zu treibenden Rechenaufwandes.

Eine weitere methodisch nicht voll geklärte Frage ist mit dem Voraufgehenden bereits angeklungen: Was ist die psychologisch zu rechtfertigende Basis für einen Vergleich der Skalen verschiedener Beurteilergruppen (für die gleichen Objekte)? Wir sind nach folgendem Grundsatz vorgegangen: Für einen Ver-

gleich sind die Skalenwerte zu wählen, welche aus der gleichen Berechnungsmethode (in dieser Untersuchung: Iteration) aufgrund des gleichen Meßmodells hervorgegangen sind, wobei eine lineare Transformation frei bleibt. Eine solche muß aber bei einer Interpretation der verschiedenen Spreizung der Skalen für alle Gruppen gleich sein. Nicht identische lineare Transformationen sind nur eingeschränkt brauchbar, etwa dann, wenn man (vgl. Abb. 6) nur einen Aspekt der Skalen verschiedener Gruppen, wie etwa die relative Veränderung von Skalenpositionen von Gruppe zu Gruppe verdeutlichen will.

Der Vergleich, den SIXTL (1967) zwischen Skalen verschiedener Berechnungsmethoden und Meßmodellen mittels einer Normierung als Vektoren im Einheitskreis durchführt, ist u. E. für unser Problem des Vergleichs zwischen verschiedenen Gruppen nicht adäquat. Denn insbesonders die dann andersartig verschiedene Spreizung ist nicht mehr psychologisch sinnvoll interpretierbar. Es wird zudem aber auch ein Kriterium des Vergleichs bei SIXTL (1967) nicht deutlich.

Weiterhin ist darauf hinzuweisen, daß bei DAWES (1977, S. 29), welcher eine Untersuchung von THURSTONE (1927) und die Daten einer Nachuntersuchung von COOMBS (1967) über die Skalierung der Schwere von Verbrechen vergleicht, die Bedeutung der verschiedenen Spreizung der Skalen nicht beachtet wird. FEGER (1972, S. 1o2) interpretiert zwar ebenfalls verschiedene Skalenspreizungen, die angewandten Transformationen und die Etablierung eines Nullpunktes bleiben aber undurchschaubar, und seine spezifischen Folgerungen sind von daher zweifelhaft. Gut begründet erscheint hingegen, bei den z-Werten als Einheit zu bleiben und im Sinne der Definition des zu skalierenden Aspektes der Krankheiten die jeweils leichteste Krankheit als "Nullpunkt" zu setzen. Das wird bei unseren Ergebnissen dadurch bestätigt, daß mit Ausnahme einer Gruppe alle anderen

der gleichen Krankheit, Grippe, den jeweils niedrigsten, zudem alle Gruppen ohne Ausnahme Krebs den höchsten Skalenwert zuordneten. Die einzige Alternative ist die von GUILFORD (1954) und SIXTL (1967) beschriebene Methode einer zusätzlichen Untersuchung zur Bestimmung eines Indifferenzpunktes einer (bzw. der) Beurteilergruppe(n). Bei der hier vorliegenden Problemstellung ist aufgrund der Definition des psychologischen Kontinuums zu erwarten, daß eben die von allen als leichteste Krankheit angesehene diesem Nullpunkt am nächsten kommt bzw. ihn am ehesten repräsentieren wird. Insgesamt gesehen ist hier also eine psychologische Lösung des Problems gelungen, im Gegensatz zu einer eher meßtechnischen. Die Absicht war von Anfang an, nicht die Relation zu einem Nullpunkt zu interpretieren, sondern durch entsprechende Wahl eines vermutlich als sehr wenig schlimm (Grippe) wie eines als sehr schlimm angesehenen Krankheitsbildes (Krebs) ein "Bezugssystem" zu erstellen, in das zumal die psychischen Erkrankungen eingeordnet werden können. Solch ein "Bezugssystem" hat den Vorteil, daß seine psychologische Bedeutung gut zu bestimmen ist, während die psychologische Interpretation von "nicht schlimm" (= o) in Bezug auf eine Krankheit große Schwierigkeiten bietet.

5.1.2 Kritische Bewertung der Ergebnisse im Licht des Untersuchungsplanes

In gleicher Weise, wie das Meßmodell und die mathematisch-statistischen Methoden im engeren Sinne ein notwendiges logisches Gerüst sind, auf welchem die Gültigkeit der Ergebnisse aus der Anwendung des Hauptuntersuchungsinstruments aufbaut, sind die Ergebnisse auch in Hinsicht auf die Gesamtanlage der Untersuchung kritisch zu bewerten, wenn man nicht einfach in ihrer allgemeinen Bedeutung nicht interpretierbare Kennwerte aufzählen will.

Diese Untersuchung will nicht nur mit einer möglichst exakten Methodik subjektive Bewertung von Krankheiten, vor allem von psychiatrischen Erkrankungen, messen, sondern auch durch den Vergleich verschiedener Gruppen, welche sich nach definierten Merkmalen unterscheiden sowie durch die Kontrolle von möglichen Einflußgrößen eine Klärung der Bedingungsfaktoren solcher Bewertungen versuchen.

Ein Kontrollinstrument, welches durch die Paarvergleichsmethode ermöglicht wird, ist die Prüfung der individuellen Urteilskonsistenz. Sie erlaubt Beurteiler zu erkennen, welche nicht hinreichend adäquat die gestellte Aufgabe des Vergleichs angehen können.

Zusätzlich wurden bei allen Gruppen Kennwerte über die schulische und berufliche Ausbildung und damit indirekt (und global, aber für diese Fragestellung bei der zusätzlichen Prüfung der Kompetenz hinreichend) über das allgemeine Intelligenzniveau erhoben und zu den Ergebnissen in Beziehung gesetzt.

Die Ergebnisse zeigen insgesamt, daß diese Faktoren nur einen indirekten Einfluß auf die Beurteilung haben. Die wesentliche und direkte Einflußgröße ist der Grad der Informiertheit über die (anders: der Grad der Kenntnis der) Krankheiten. Die anderen Kennwerte repräsentieren hingegen Faktoren, welche die Wirkung des wesentlichen Faktors sozusagen limitieren.

Der Grad der Informiertheit über Krankheiten wurde bereits im voraus als Einflußfaktor von erheblicher Bedeutung vermutet.

Daher wurde versucht, diesen Faktor einmal mit Hilfe des Wissenstests zu kontrollieren (bei Kontr 1, Tb und Alk). Zum zweiten wurde versucht, Informationsgrad allgemein (Kontr 2) und spezielle Information über psychiatrische Erkrankungen als gestufte Einflußvariable einzuführen (Kontr 2, Med 1 zu Med 2).

Vor allem aus dem Vergleich von Med 1 und Med 2 ergibt sich mit hoher Wahrscheinlichkeit, daß ein erhöhtes Maß an differenzierter Information die Bewertung psychiatrischer Krankheiten verändert. Eine methodische Unzulänglichkeit der Untersuchung wird jedoch zugleich deutlich: Die Veränderung läßt sich nicht sicher als alleiniger Effekt der Information ausmachen. Von den Ergebnissen her wird das dadurch deutlich, daß Med 2 weniger zwischen den Krankheitsbildern differenziert als Med 1. Es gibt dafür verschiedene mögliche Gründe: Einmal könnte durch die Wiederholung ein Sättigungs- oder "Überdruß"-Effekt eingetreten sein. Nahegelegt wird diese Erklärung durch die Beobachtung, daß Med 2 weniger Kriterien nennt als Med 1. Gegen diese Erklärung spricht aber, daß die Häufigkeit von zirkulären Triaden in beiden Gruppen nicht signifikant differiert (Tab. 5; Chi^2-Prüfung der Häufigkeitsverteilungen gegeneinander ergab einen Kennwert von 2.91, ns). Bei einem generellen "Überdruß"-Effekt müßte ein inkonsistenteres Urteilen im Paarvergleich erwartet werden.

Die wahrscheinlichere Erklärung dürfte sein, daß der Zusammenhang zwischen Grad detaillierter Information und Differenzierung zwischen den Krankheiten (operationalisiert durch die Spreizung der Skala) nicht linear ist.

Entsprechend ist dann aber für den Unterschied zwischen Kontr 1 und Kontr 2 zu folgern, daß die Vorgabe des Wissenstests als Lerninstrument nicht so viel Information gegeben hat, daß wesentliche Veränderungen bewirkt wurden, zumindest nicht bei der Beurteilung von psychiatrischen Erkrankungen. Nur ist nun weiter zu erklären, warum die Gruppe Kontr 1 im Wissenstest signifikant besser abschneidet als Kontr 2. Hier liegt nahe, daß es eine Interaktion zwischen Wissenstest und Paarvergleichsliste gegeben hat, daß also weniger die inhaltliche Information als die gedank-

liche Beschäftigung mit den vorgestellten Krankheitsbildern ein differenzierteres Überlegen beim Paarvergleich bzw. beim Wissenstest in Gang brachte.

Obwohl der Wissenstest - wie bei der Gruppe Tb zu sehen ist - im großen und ganzen seine Funktion als Kontrolle erfüllt, wäre hiernach doch eine genauere Konstruktion im Hinblick auf eine weitere Verbesserung der Brauchbarkeit als Kontrollinstrument zu wünschen. Im Rahmen dieser Arbeit war das aus untersuchungstechnischen Gründen nicht erreichbar. Wir nehmen aber an, daß dann die korrelativen Zusammenhänge zwischen Bildungsstand, Anzahl zirkulärer Triaden und Gesamt-Rohwert im Wissenstest (vielleicht schon aus Gründen der stärkeren Streuung im Wissenstest) deutlicher hervorgetreten wären.

Ähnlich begrenzt durch Bedingungen der Untersuchungsdurchführung war die Möglichkeit, die Kriterien der Beurteilung detailliert zu erfassen. Hinreichend belegt wurde, daß alle Gruppen in etwa nach den gleichen, am konventionellen Bild einer körperlichen Krankheit orientierten Vorstellungen urteilten. Eine differenziertere Methode zur Erfassung von Urteilskriterien mittels einer Nachbefragung und Interpretation der dann vorliegenden Skala durch die Gruppe selber wies SÜLLWOLD (1969 b) auf.

Den Einfluß der Variable Geschlecht kann man durch die Teilung der Gruppen Med 1 und Alk grob abschätzen: Bei beiden weiblichen Gruppen zeigte sich, daß psychiatrische Erkrankungen schwerer bewertet wurden. Es ist aber nicht möglich, Schlußfolgerungen zu ziehen, warum diese Unterschiede bestehen. Wünschenswert wäre eine - praktisch aber nicht erreichbare - Gruppe von Beurteilerinnen gewesen, welche den Gruppen Kontr 1 und Kontr 2 im Hinblick auf Alter und Bildungsstand vergleichbar wäre.

Während alle aufgeführten Einflußvariablen zumindest

annähernd abgegrenzt werden konnten, so daß wenigstens eine Hypothesenbildung über ihre Wertigkeit möglich war, konnte das nicht mit der Variable Alter geschehen. Von ihr ist zu vermuten, daß ihr Einfluß erheblich ist. Die Gruppe Tb weist die größte Altersstreuung auf, zugleich den größten Anteil an älteren Beurteilern. Es ist daher noch einmal darauf hinzuweisen, daß hier die in manchem deutlich abweichende Skalierung der Krankheiten nicht nur auf die aufgezeigten kontrollierten Variablen zurückzuführen sein dürfte, sondern daß deren Einfluß vermutlich wesentlich durch den Faktor Alter konfundiert ist.

Bisher sind die für eine Gesamtbetrachtung kritischen Punkte aufgeführt worden, welche sich immanent aus dem unvermeidlich nicht perfekten Untersuchungsplan ergeben. Der Vergleich mit bisherigen und von uns referierten Untersuchungen zeigt jedoch noch weitere einschränkende Gesichtspunkte auf. Sie wurden aber bewußt gewählt, sind bei einer Bewertung des Gesamtergebnisses jedoch zu bedenken.

Wir haben von vornherein beabsichtigt, in der Instruktion ein möglichst eindimensionales Kriterium zu formulieren. Diese Vorentscheidung ist durch die dem Meßmodell inhärenten Prüfmöglichkeiten bestätigt worden. Aber damit ist nur belegt, daß sich eine vorherrschende Beurteilungsdimension aufzeigen läßt, nicht aber die Unmöglichkeit eines mehrdimensionalen Modells. Denn die Wahl einer Formulierung, welche versucht, ein möglichst globales Kriterium zu wählen, "für wie schlimm hielten Sie, wenn ..." geschah deswegen, weil vermutet wurde, daß etwa ein betroffener Patient sich ein persönliches, summarisches Gesamturteil über seine Situation bildet. Wir meinen, daß dies für ihn wesentlicher ist als sicher auch vorhandene, aber nur partielle Aspekte betreffende Erwägungen im Blick auf etwa verschiedene Auswirkungen der

Krankheit, z. B. im Hinblick auf den Beruf, die Familie, die Freizeit, Sozialkontakte etc. Die Verbindung zum ersten Gesichtspunkt, der psychologischen Eindimensionalität, ist darin zu sehen, daß bei Vorherrschen solcher Einzelaspekte eben keine eindimensionale Skalierbarkeit zu erwarten gewesen wäre.

Dagegen geben die von uns referierten Untersuchungen zumeist das Kriterium "Angst" vor. Über diese Untersuchungen hinweg ergeben sich Unterschiede deswegen, weil der Grad der Angst offensichtlich moderiert wird durch das, was JENKINS & ZYZANSKI (1968) als "saliency" bezeichnen. Es meint sozusagen ein persönliches Urteil über die Wahrscheinlichkeit, selber betroffen zu werden, verbunden mit daraus folgender mehr oder weniger häufiger gedanklicher Beschäftigung mit einer Krankheit. Bei allen genannten Untersuchungen ist dieser moderierende Faktor für "Grad der Angst" in seinem Gewicht nicht hinreichend einzuschätzen, ausgenommen wohl bei Krebs, wo sich aufgrund einer häufigen Einwirkung der Medien allgemein hohe "saliency" vermuten läßt. In unserer Untersuchung versuchten wir auch deswegen, die Situation eines Betroffenen in der Vorstellung der Beurteiler (ausgenommen bei Med 1 und Med 2) zu reproduzieren, einmal durch die explizite Instruktion "selber betroffen", zum anderen durch die Einbeziehung von Patientengruppen. Man kann aufgrund der Ergebnisse davon ausgehen, daß Unterschiede in der persönlichen Beurteilung der "Wahrscheinlichkeit" dadurch ausgeglichen wurden.

Auf einem anderen Hintergrund zu sehen sind die Kontraste zu der Untersuchung von BINDER & ANGST (1981). Sie fragen nach "Beschwerden". Ähnlich geht STUMME (1975) vor, welcher Eigenschaften skalieren läßt, die sozusagen nach Meinung von Laien "pathognomonisch" sind für psychische Krankheit. In der vorliegenden Untersuchung sind die vorgegebenen Begriffe in der Instruk-

tion hingegen klar als "Krankheiten" definiert. Diese Definition ist sicher zusammen mit dem Kontext und der Vorstellung als Substantiva eine Vorausfestlegung für die Einordnung der psychischen Bilder in die Reihe der organischen Krankheiten. Konkreter gesprochen: Eine Überlegung, ob insbesondere Neurose und endogene (das Wort war zumeist erklärungsbedürftig) Depression wirklich "richtige" Krankheiten seien mit möglichen Auswirkungen auf die Beurteilung des "Schlimmheits"-Grades wurde damit präjudiziert, und zugleich wurde wohl die psychologische Eindimensionalität verstärkt. Aus der Untersuchung von BINDER & ANGST (1981) war referiert worden, daß dort ein deutlicher Unterschied zwischen Grad subjektiver Belastung (höher bei psychischen Beschwerden) festgestellt wurde. In der vorliegenden Untersuchung ergibt sich eine Parallele insofern, als bei der Befragung nach den Kriterien festgestellt werden konnte, daß in allen Gruppen diese weitgehend "körperlich" akzentuiert sind, daß in allen befragten Gruppen der Nichtmediziner etwa die sozialen Konsequenzen gar nicht erwähnt werden, ausgenommen - nach Anreiz - bei der Gruppe Tb, die aber sehr persönliche Erfahrungen reflektiert, ohne daß diese sich auf die vorhergegangene Einschätzung der eigenen Krankheit im Kontext der anderen gegenüber den Vergleichsgruppen auswirkt. Und auch bei Krankheiten, die den Aspekt sozialer Belastung haben könnten, etwa eben Tuberkulose, aber auch Akne als Hautkrankheit und Syphilis als Geschlechtskrankheit, läßt sich in diesem Zusammenhang eine besondere Bedeutung sozialer Gesichtspunkte nicht ausmachen. Jedoch wäre natürlich die Folgerung nicht berechtigt, daß spezifische Belastungen aus dem sozialen Umfeld keine Rolle spielen, insbesondere für das individuelle Erleben eines faktisch Betroffenen. Nur scheinen diese Aspekte keine spontan auftauchenden Bewertungsgesichtspunkte zu sein und werden von den Experten wahrscheinlich überbewertet.

5.2 Die Urteilsunsicherheit über die psychiatrischen Erkrankungen: Überlegungen zu psychologischen und therapeutischen Konsequenzen

Bei den methoden- und untersuchungskritischen Überlegungen, denen ein großes Gewicht beizumessen ist, soll jedoch nicht übersehen werden, daß generell eine subjektive Skala der Schwere von Krankheiten festgestellt werden konnte. Über sie besteht in nach Vorkenntnissen, Bildungsniveau, persönlicher Erfahrung und - sekundär eruiert - Geschlecht sowie nach Alter differierenden Gruppen eine sehr große Übereinstimmung.

Krankheiten wie Grippe, Akne und Blinddarmentzündung werden als wenig schlimm angesehen; z. T. auch noch Bluthochdruck. Zu den schlimmeren Erkrankungen wird Zuckerkrankheit gezählt, wechselweise "von unten" sozusagen auch Bluthochdruck, "von oben" Syphilis. Als noch schlimmer bewertet lassen sich Syphilis, Tuberkulose und Herzinfarkt in eine weitere Gruppe zusammenfassen, als mit Abstand schlimmste gilt Krebs, als zweitschlimmste Querschnittslähmung. Es ist aber ebenso möglich, statt einer Vierergruppierung eine Dreiergruppierung vorzunehmen, wobei dann die beiden mittleren Kategorien in eine zusammenfallen.

Die Übereinstimmung zwischen den Beurteilergruppen geht weitgehend verloren im Hinblick auf ihre Beurteilung psychischer bzw. psychiatrischer Erkrankungen. Als gemeinsamer Rahmen bleibt, daß sie weder als leicht noch als sehr schlimm beurteilt werden, und daß Neurose übereinstimmend als am wenigsten schlimm beurteilt wird. Untereinander und in Relation zu den körperlichen Erkrankungen wechseln sie jedoch die Skalenpositionen über die Gruppen hinweg, auch bei einer sekundären Unterteilung geeigneter Beurteilergruppen nach Geschlecht, wobei weibliche Beurteiler sie

generell als schlimmer ansehen.

Es folgt also, daß die Einflußgrößen, die in den verschiedenen Gruppen unterschiedlich wirksam sind, sich fast ausschließlich auf die Bewertung psychischer Krankheiten auswirken. Vor allem der unterschiedliche Grad der Informiertheit ist von großer Bedeutung, und damit indirekt der Bildungsstand. Weiterhin zeigt sich ein deutlicher Geschlechtsunterschied, und aller Wahrscheinlichkeit nach - leider aus Stichprobengründen nicht eindeutig zu belegen - ein Alterseinfluß. Persönliche Erfahrung mit einer Krankheit verändert dagegen in den Gruppen der Tuberkulose- und Alkoholpatienten die Skalenposition der eigenen Krankheit nicht (in Relation zu den anderen Gruppen).

Die Beurteilungskriterien sind im wesentlichen "konventionell" geprägt - wie wohl auch das Bild von der Medizin: Körperliche Beeinträchtigungen, Lebensgefahr, Heilbarkeit sind die wesentlichen Gesichtspunkte in allen Gruppen. Soziale Folgen oder soziales Image einer Krankheit dürften dagegen keine merkbare Wirkung auf die Beurteilung haben. Das gilt wohl auch für die Beurteilung der psychiatrischen Krankheiten. Nur die Mediziner setzen Schizophrenie deutlich von Hirnarteriosklerose ab, während bei den meisten anderen Gruppen die "organische" Konnotation von "Hirn"-arteriosklerose durchschlagen dürfte.

Es läßt sich weiterhin wahrscheinlich machen, daß eine Steigerung der Information die Bewertung ändert, weniger bei den körperlichen Erkrankungen, deutlich bei den psychiatrischen Krankheitsbildern. Allgemein scheint sich ein höherer Grad von Information vor allem im Hinblick auf die Differenzierung zwischen den Krankheiten auszuwirken, aber es gibt starke Hinweise darauf, daß diese Beziehung nicht linear ist. Als Hypothese ließe sich formulieren, daß geringere Kenntnis weniger gut unterscheiden läßt, größere Kenntnis

zu besserer Differenzierung führt, noch komplexere Information aber die Unterscheidung wieder schwieriger macht und vielleicht dann weitere Urteilsdimensionen ins Spiel bringt. Eine hypothetische Reihenfolge des Informationsgrades in Bezug auf die Gruppen wäre: Tb, Alk, Kontr 1, Kontr 2, Med 1, Med 2.

Es darf aber nicht außer Acht gelassen werden, daß sich "objektiv" die einzelnen Krankheitsbezeichnungen ihrem Inhalt nach wesentlich unterscheiden, was vor allem den angehenden Ärzten bei steigender Information deutlicher werden dürfte. Es gibt geschlossenere und wenig geschlossene Einheiten. Das Ergebnis der guten linearen Skalierbarkeit darf jedoch als Stütze dafür betrachtet werden, daß vor allem die für den Experten sehr globalen und bis zu einem gewissen Grad sogar Heterogenes beinhaltenden Bezeichnungen für Gruppen aus der Durchschnittsbevölkerung inhaltlich sinnvoll, und wie zumal der Vergleich zwischen den beiden Kontrollgruppen mit der Variation der Abfolge der Paarvergleichsliste und Wissenstest zeigt, auch hinreichend zutreffend gefüllt sind.

Durch die Betrachtung von kontrastierenden Gruppen, welche sich nicht in erster Linie nach Randdaten, sondern nach der Einstufung der Krankheiten selber unterscheiden, wird die Bedeutung des Informationsstandes unterstrichen, und zwar spezifisch im Hinblick auf die Einordnung psychischer Erkrankungen.

Es war zu sehen, daß eine in Relation zu den ursprünglichen Gruppen (Tb, Alk, Kontr 1 und Kontr 2) extrem niedrige Einschätzung psychischer Krankheiten einen deutlichen Zusammenhang mit geringerer Kenntnis über diese Krankheitsbilder, andeutungsweise über alle Krankheitsbilder, mit geringerem Bildungsstand allgemein zusammenhängt. Schaltet man diesen Faktor durch Herausnahme entsprechender Beurteiler (nach den Daten des "Wissenstests") aus, lassen sich im strengen

Sinn Kontrastgruppen, die von den Gesamtgruppen in der Bewertung psychischer Krankheiten nach unten abweichen, nicht mehr identifizieren. Daraus ist zu folgern, daß eine dezidierte Einstellung, welche psychische Beeinträchtigung bzw. Erkrankung leicht nimmt, zumindest selten ist (bei den hier befragten Beurteilern!), so daß sie das Gruppenurteil nicht wesentlich mitbestimmt.

Zugleich lassen sich aber Einzelfälle finden - ein Beispiel ist angeführt -, die durchaus aus einer Einstellung heraus urteilen in dem Sinne, daß wirklich schwerwiegend körperliche Faktoren sind, während Psychisches eher als leicht angesehen wird, sie zeigen sich dabei zugleich als hinreichend gut informiert.

Eine "starke Minderheit" kontrastiert hingegen in die andere Richtung: Der Faktor des Psychischen macht eine Krankheit schlimmer, eine körperliche Beeinträchtigung oder Gefährdung wird hingegen weniger schlimm gesehen. Diese Gruppe unterscheidet aber nicht nach einzelnen seelischen Erkrankungen, sondern stuft alle relativ einheitlich bei den schweren Krankheiten ein, so wie es auch insgesamt für die Ursprungsgruppen zu erkennen war.

Welche Konsequenzen aus diesen Ergebnissen könnten zumal für die therapeutische Arbeit mit psychiatrischen Patienten abgeleitet werden? Eine erste Überlegung ist, daß man Einschätzungen und Meinungen zu Krankheiten, die auch bei psychischen Erkrankungen vorhanden sind, wenn auch die Urteilsunsicherheit größer ist, nicht einfach bei der Behandlung übergehen kann. Je schlimmer die Erkrankung eingeschätzt wird, desto schwieriger wird das sein. Von sehr großer Wichtigkeit ist das aber gerade bei psychischen Erkrankungen aufgrund der hier besonders starken Unsicherheit.

Die zweite Konsequenz bezieht sich auf die Ausgangsfrage der hier vorgelegten Untersuchung. Danach er-

scheint es sinnvoll, den Effekt einer eingehenden Besprechung und Erklärung einer psychiatrischen (insbesonderes psychotischen) Erkrankung als Grundlegung einer üblicherweise langdauernden therapeutischen Betreuung in weiteren Untersuchungen empirisch abzuklären.

Es ist immerhin begründet zu vermuten, daß damit eine bessere Einsicht des Patienten in seine Situation und in die Möglichkeiten seiner eigenen Mitarbeit bei der Therapie erreicht wird, was die nachgewiesene Wirksamkeit etwa verhaltenstherapeutischer Trainingsprogramme erhöhen könnte. Dabei ist hier natürlich nicht an Zeiten noch produktiv-psychotischer Krankheitsstadien gedacht. Sinnvollerweise sollte ein solcher Versuch alsbald einsetzen, wenn produktive Symptome zurücktreten. Realistisch muß man zugleich aber annehmen, daß eine solche Möglichkeit nicht bei allen Patienten gegeben ist; die Bedeutung etwa des Grades der Störung wie auch etwa der prämorbiden intellektuellen Befähigung ist sicherlich abzuklären. Offen ist das Problem, wie man die Mitteilung einer Diagnose formulieren soll.

In den bisher formulierten Überlegungen sind eher weitere Forschungsfragen als gesicherte empirische Erkenntnisse enthalten. Als sichere Folgerung aus den Ergebnissen der Untersuchung sollen aber zwei Punkte herausgestellt werden:
Zum einen konnte gezeigt werden, daß trotz der erheblichen Urteilsunsicherheit über die genaue Bedeutung psychiatrischer bzw. psychischer Erkrankungen die Begrifflichkeit bekannt ist und auch eine im groben zutreffende inhaltliche Vorstellung, z.B. von "Schizophrenie" gegeben ist. Von daher ist es notwendig, solche Begriffe auch aufzunehmen und zu besprechen. Wie eigene noch im exploratorischen Stadium befindliche Versuche zur begleitenden psychotherapeutischen Behandlung von schizophrenen Erkrankungen ergaben, kannte auch jeder der Patienten diesen Begriff. Jeder

wußte (ohne daß die behandelnden Ärzte diesen Terminus gebraucht hatten), daß die eigene Erkrankung so bezeichnet wurde, und daß etwa in Lexika Schizophrenie als "Geisteskrankheit" definiert wird. Nach bisherigen Erfahrungen ist es aber ebenso notwendig, dann nicht bei diesem "objektiv" globalen und zu undifferenzierten Begriff stehen zu bleiben. Bisher erscheint es fruchtbar, die Bedeutung der psychiatrischen Terminologie zu erklären, weiterhin eingehend darzustellen, wie heute die Erkenntnisse über die Erkrankung fortgeschritten sind, aber auch, was man noch nicht hinreichend genau weiß (z. B. mangelnde Kenntnis des genauen Wirkmechanismus der Medikamente, was auf das Problem der Nebenwirkungen und auf die Notwendigkeit eines oft langwierigen individuellen Erprobens der optimalen Dauermedikation führt).

Zum zweiten konnte nachgewiesen werden, daß bei den Kriterien für die Beurteilung der subjektiven Schwere einer Erkrankung die Überzeugung über die "Heilbarkeit" von wesentlicher Bedeutung war. Im konkreten Fall (nicht nur natürlich bei psychischen Krankheiten) fragt sich jeder betroffene Patient in Bezug auf seine eigene Erkrankung danach. Er stuft danach auch ein, wie schwer er erkrankt ist, verbindet vor allem seine Befürchtungen und Ängste damit. Aufgrund der bei der Mehrheit weniger dezidierten Vorstellungen über psychische Erkrankungen im Gegensatz zu körperlichen Krankheiten schwanken diese Einschätzungen stärker. Auch diese Vorstellungen sind mit Sicherheit aufzugreifen, wobei schon fast trivial ist anzumerken, daß nach den Ergebnissen der sozialpsychologischen Forschung über Einstellungs- bzw. Meinungsänderungen eine zu große Abweichung von der Überzeugung dessen, den man zu beeinflussen sucht - in diesem Falle also vor allem ein Versuch der Bagatellisierung - negative Effekte zeitigen wird.

Eine abschließende Überlegung schließlich soll den von

uns herausgestellten Aspekt der Urteilsunsicherheit mit der referierten Untersuchung von KNOPF (1976) verbinden, welche auf den Prozeß der öffentlichen Meinungsbildung über Vorsorge und Behandlung einmal von Krebserkrankungen, zum anderen von psychischen Erkrankungen in Großbritannien abhebt. Die "Fremddarstellung" der Psychiatrie, aber - da hier in der Öffentlichkeit kaum unterschieden wird - wohl auch der Psychologie als Klinischer Psychologie als auch psychologischer Therapie im deutschsprachigen Bereich ist sicher in gleicher Weise negativ zu charakterisieren.

Wünschenswert erscheinen eingehendere Kenntnisse darüber, ob nicht eine in dieser Weise verstärkte Unsicherheit vor allem Betroffene oder potentiell Betroffene sehr stark beeinflußt und therapeutische Bemühungen konterkarieren kann. Das erscheint immerhin denkbar.

Es stellt sich aber auch die Frage nach der Möglichkeit einer Meinungsbildung über psychische Erkrankungen in Kommunikationsmedien. Bei einer Selbstdarstellung von Psychiatrie und Psychotherapie wird üblicherweise betont, wie groß der Bedarf an Verbesserungen ist. Dabei pflegt man aber auch immer zu vermitteln, wie schwer die Krankheiten sind, die hier behandelt werden müssen. Man entspricht damit der Einstellung einer abgehobenen Gruppe, wie aus den Ergebnissen des hier vorgestellten Kontrastgruppenvergleichs gefolgert werden kann. Diese abgehobene Gruppe teilt die Einstellung, daß psychische Erkrankungen zu den schlimmsten Krankheiten gehören. Die Mehrheit, so ist zu vermuten, setzt jedoch andere Prioritäten, wobei sie durchaus nicht bagatellisierend eingestellt ist, wenn auch im Urteil unsicherer ist. Daß diese Folgerung eine Generalisierung aus den hier vorgelegten Ergebnissen ist, welche das Problem der Repräsentativität der Beurteilergruppen berührt, muß dabei betont werden. Wahr-

scheinlich ist jedoch, daß eine solche Generalisierung für jüngere, vielleicht auch für mittlere Altersgruppen möglich ist. Unter diesen Voraussetzungen ist aber dann möglicherweise vieles an öffentlicher Selbstdarstellung und Diskussion von seiten der Psychiatrie und auch der Psychologie die Kommunikation einer gebildeten Minderheit mit sich selber, erreicht aber nicht für die Problemlage durchaus aufnahmebereite und ansprechbare Außenstehende.

6 Zusammenfassende Darstellung

Es sollen in einem Überblick noch einmal die wesentlichen Ausgangspunkte dieser Untersuchung, die entscheidenden Aspekte zur Untersuchungsmethodik und die Ergebnisse dargestellt werden.

6.1 Hintergrund und empirisch bearbeitetes Umfeld

Klinische Erfahrungen mit langfristiger Betreuung psychiatrischer Patienten zeigen, daß das Anfangsstadium einer psychischen Erkrankung für eine früh einsetzende Mitarbeit eines Patienten an seiner Behandlung ein kritischer Zeitraum ist.

Es gibt vielfältige gesicherte Arbeiten zur Pharmakotherapie insbesondere der Psychosen, dann auch eine große Zahl von Untersuchungen zur ergänzenden psychologischen Therapie. Für das Anfangsstadium vieler psychischer Krankheiten besteht aber eine empirische Lücke an gesicherten Erkenntnissen über das Denken und Erleben wie auch die mitgebrachten Einstellungen und Meinungen der Betroffenen. Dem entspricht eine weit verbreitete Unsicherheit auf seiten der Behandelnden, die zumal dann deutlich wird, wenn ein an einer schizophrenen Psychose Erkrankter nach seiner Diagnose fragt.

Das bisher empirisch bearbeitete Umfeld der vorliegenden Untersuchung läßt sich in aller Kürze kennzeichnen durch die Nennung von drei Themenkreisen:

- Einstellung zum psychisch Kranken, einschließlich der Untersuchungen zum Autoheterostereotyp des psychisch Kranken selber (JAECKEL & WIESER 197o, STUMME 1975, BINDER & ANGST 1981, VAN QUEKELBERGHE 1976)
- Compliance bei psychiatrischen Erkrankungen (HERMANN 1979, LINDEN 1979)
- vor allem anglo-amerikanische Untersuchungen über

Einstellungen zu Krankheiten, wobei "mental illness" aber nur nebenbei behandelt wird (ROBBINS 1962, JENKINS & ZYZANSKI 1968, LEVINE 1962, KNOPF 1976).

6.2 Fragestellung

Das Interesse dieser Arbeit richtete sich auf folgende grundlegende Fragen für den Bereich "Einstellung zu und Meinungen über psychische Krankheiten":

1. Verbinden Gruppen aus der Bevölkerung überhaupt eine hinreichend einheitliche und hinreichend richtige Vorstellung mit bestimmten Krankheitsbezeichnungen, vor allem mit psychiatrischen Krankheiten (Definiertheit des psychologischen Objekts)?
2. Wenn man Gruppen bittet, Krankheiten danach zu skalieren, für wie schlimm sie diese halten, wenn sie selber betroffen wären: Wo rangieren dabei die psychiatrischen Krankheiten (Einstellung zur Krankheit unter einem definierten Aspekt)?
3. Skalieren Gruppen, welche eigene Erfahrung mit einer Krankheit haben, anders als Gruppen ohne konkrete Krankheitserfahrung?
4. Lassen unterschiedliche Grade des Informationsstandes solche Skalen unterschiedlich aussehen?
5. Unterscheiden sich einzelne Beurteiler oder Gruppen von Beurteilern deutlich in der Einstufung von psychiatrischen Krankheiten als entweder sehr schlimm oder im Gegensatz dazu wenig schlimm? Finden sich dann mit einer solchen kontrastierenden Beurteilung korrespondierende Unterschiede in anderen Merkmalen, welche Hinweise auf die Gründe solch unterschiedlicher Beurteilung geben können?

6.3 Entwicklung der Untersuchungs- und Auswertungsmethodik

Die Betrachtung der wenigen bisherigen Untersuchungen zur Frage der Einschätzung von Krankheiten nach verschiedenen Beurteilungsdimensionen ergab, daß sowohl

der Methode der Datengewinnung wie auch der statistischen Auswertung eingehende Überlegungen zu widmen waren.

Als Skalierungstechnik wurde - zuerst unter einem pragmatischen Gesichtspunkt - die Paarvergleichsmethode nach THURSTONE (1927) gewählt. Dabei hat der einzelne Proband jeweils alle möglichen paarweisen Kombinationen der zu beurteilenden Objekte zu betrachten. Er muß jeweils entscheiden, welcher von beiden Paarlingen die definierte Eigenschaft in stärkerer Ausprägung repräsentiert. Für jeden einzelnen Beurteiler resultiert daraus eine sogenannte Dominanzmatrix und aus dieser eine Rangreihe der Objekte. Die Dominanzmatrizen der einzelnen Beurteiler lassen sich nach definierten Regeln zu einer Gesamt-Präferenzmatrix einer Gruppe zusammenfassen, aus welcher dann eine Skala der Objekte berechnet werden kann. Diese Methode hat deswegen Vorzüge, weil sie von den Beurteilern eine einfache, überschaubare Operation verlangt, weil außerdem alle Voraussetzungen des zugrunde gelegten Meßmodells aus den Daten empirisch prüfbar sind.

Diese Voraussetzungen, welche die Datenstruktur erfüllen muß, um das für die Interpretation erwünschte Intervallskalenniveau zu erreichen, wurden diskutiert. Es konnte gezeigt werden, daß die meßtheoretischen Axiome der Konnexität (ORTH 1974, GIGERENZER 1981), d. h. vollständige Ordnung aller Objekte durch die Beurteiler, und vor allem der Transitivität erfüllt waren. Deutlich wurde, daß die relativ umständliche Paarvergleichsoperation (bei insgesamt 15 Krankheiten sind 1o5 Paarkombinationen zu beurteilen) tatsächlich den Vorteil bietet, daß vor allem die Erfülltheit der Transitivitätsforderung exakt geprüft werden kann. Damit konnte in einem ersten Schritt belegt werden, daß jeder Beurteiler tatsächlich auf einem ordinalen Skalenniveau urteilt, bzw. hiervon abweichende Beurteiler konnten identifiziert und ausgeschieden werden. Das war

die hinreichende Voraussetzung zur Entwicklung jeweils einer gemeinsamen Skala für jede nach Gesichtspunkten der Homogenität zusammengestellte Beurteilergruppe, welchen Skalen man dann Intervallskalenqualität zusprechen kann.

Im Sinne der Fragestellung war ein reizzentriertes Skalierungsverfahren - die Paarvergleichsmethode zählt dazu - gewählt worden. Dem entgegen steht das von GIGERENZER (1981) aus modelltheoretischen Überlegungen abgeleitete Verdikt, eine reizzentrierte Skalierungsmethode sei prinzipiell einem psychologischen Forschungsgegenstand nicht angemessen. Sie wende ein aus der klassischen Physik stammendes zweistelliges Modell (Objekt und Eigenschaftsausprägung am Objekt) auf einen sozialwissenschaftlichen Tatbestand an, klammere damit die gerade für die Psychologie wesentliche Individualität der Beurteiler, welche sich in der Variation der Urteile niederschlage, aus und mache diese zu bloßen "Meßinstrumenten". Eine Analyse der von GIGERENZER (1981) postulierten fünfstelligen Modellrelation konnte aufweisen, daß eine methodisch notwendige Ausblendung der interindividuellen Variation zwischen den Beurteilern nicht gleichzusetzen ist mit einer prinzipiellen Negierung ihrer Bedeutsamkeit. Zugleich wurden die Überlegungen von SIXTL (1972) aufgenommen, welcher in einem ersten Ansatz empirische Kriterien für die Legitimität der Bestimmung eines durchschnittlichen Urteils einer Gruppe bei einer reizzentrierten Skalierung aufgestellt hat. Es konnte belegt werden, daß ein solches Vorgehen legitim ist bei homogenen Gruppen, welche im wesentlichen nur zufällige Urteilsabweichungen voneinander bieten. Zugleich ergab sich, daß Bedingungsfaktoren für potentiell "definitive" Urteilsunterschiede bereits durch das Untersuchungsdesign analysierbar gemacht werden müssen.

Es erwies sich daher als notwendig, die Beurteilungsprozedur nicht zu verkürzen (etwa im Sinne einer direk-

ten Skalierung) und von daher auf eine umfangreiche Erhebung zusätzlicher Daten zu verzichten. Für die freiwillige Mitarbeit der verschiedenen Beurteilergruppen war ein begrenzter Zeitrahmen vorgegeben, der durch die Bearbeitung der Paarvergleichsliste von 15 Krankheiten und die Beantwortung eines Wissenstests über diese Krankheiten gefüllt war.

Bei der Verwendung statistischer bzw. rechentechnischer Einzelverfahren konnte im wesentlichen auf die Gesamtdarstellung von SIXTL (1967) für dieses Skalierungsverfahren aufgebaut werden. Allerdings boten sich bei einzelnen Verfahrensteilen noch ungelöste Probleme, deren Behandlung von Bedeutung für die Interpretation der Ergebnisse war. Am Beispiel einer Beurteilergruppe wurde das mathematische Verfahren erläutert, und zugleich wurden die eigenen Lösungen für Einzelprobleme dargestellt.

So war im Hinblick auf die Transitivitätsforderung ein neues Kriterium für das zu tolerierende Ausmaß an Inkonsistenz des Urteilens (Anzahl sog. "zirkulärer Triaden" pro Beurteiler) zu entwickeln. Das Kriterium nach KENDALL (1948) erwies sich insofern als unzureichend, als sich ein generelles zufallskritisch gesichertes Maß nicht finden läßt. KENDALLs Annahme, daß es für alle Beurteiler eine gleiche maximal mögliche Anzahl von zirkulären Triaden gibt, trifft nicht zu. Vielmehr hängt diese maximal mögliche Zahl aus logischen Gründen vom jeweiligen konkreten Urteilsverhalten des einzelnen ab (SIXTL 1967). Es wurden deswegen ein arbiträres, gegenüber der KENDALL-Methode aber für alle Fälle strengeres Kriterium gewählt.

Als Bedingungsfaktoren für intransitives Urteilen (also Zustandekommen von zirkulären Triaden) kommen in Frage mangelnde Kompetenz der Beurteiler, mangelnde Eindimensionalität des definierten Beurteilungskriteriums und geringe "objektive" Differenz der Objekte auf

der zu erschließenden Skala. Mangelnde Eindimensionalität konnte für alle Gruppen ausgeschlossen werden. Eine numerische Bestimmung der Zusammenhänge zwischen Bildungsstand (kombinierter Index aus Angaben über Schul- und Berufsausbildung), Kenntnis der Krankheiten (Wissenstest) und Anzahl tatsächlicher zirkulärer Triaden (KENDALL-Formel) über den Eta-Koeffizienten (andere Zusammenhangsmaße konnten aufgrund der differierenden Verteilungsformen der Kenngrößen nicht herangezogen werden) belegte weiter für alle Gruppen, daß mangelnde Beurteilerkompetenz nur schwach korreliert war mit der Anzahl der zirkulären Triaden. So ergab sich, daß vor allem der dritte Faktor, geringe "objektive" Differenz auf der Skala, bestimmend war.

Unter den möglichen Alternativen für die mathematische Umwandlung der aus den Rohdaten resultierenden Präferenzmatrizen war die Wahl zwischen den von THURSTONE (1927) formulierten Varianten des "Law of Comparative Judgement" (Fall V und Fall III) mit den sie unterscheidenden Annahmen über die Art der Streuung des diskriminativen Verhaltens zu diskutieren. Dabei beruht die Rechtfertigung der Annahme der Intervallqualität der resultierenden Gruppenskalen auf dem empirisch nachweisbaren Faktum einer Urteilsstreuung. Es zeigte sich bei dieser Untersuchung, daß die im Sinne der Fragestellung notwendige weitgestreute Auswahl der Objekte (Krankheiten) in hoher Zahl sog. "Extremverteilungen" der Präferenzhäufigkeiten provozierte, d. h. eine Krankheit eines Paares wird von allen Beurteilern einer Gruppe als schwerer angesehen (relative Häufigkeit der Bevorzugung von 1oo % bzw. 1.o). Aus diesem Grund erwies sich das Iterationsverfahren nach GULLIKSEN (1956) als das angemessenste Verfahren, die Gruppenskalen aus den Präferenzmatrizen zu bestimmen (mathematisch: die jeweilige Matrix in einen Vektor umzurechnen). Zugleich war mit diesem Verfahren überflüssig geworden, bei Extremverteilungen willkürliche z-Werte einzusetzen, weil es auch die Behandlung

unvollständiger Präferenzmatrizen gestattet.

Die Prüfung der Güte der Anpassung des Vektors an die Ausgangsmatrix wird üblicherweise so vorgenommen, daß aus diesem Vektor wiederum eine Matrix zurückgerechnet wird. Es zeigte sich, daß das Chi^2-Verfahren nach MOSTELLER (1951), mit dem die beiden Matrizen auf Übereinstimmung geprüft werden, dem Iterationsverfahren nicht angemessen ist. So wurde die Bestimmung der Indices der (linearen) Skalierbarkeit nach GULLIKSEN & TUKEY (1958) gewählt. Hier wird nicht die Güte der Anpassung beider Matrizen geprüft - es konnte gezeigt werden, daß das beim Iterationsverfahren auch redundant ist -, sondern ähnlich einem varianzanalytischen Modell der Anteil der Matrixvarianz bestimmt, die linear erklärbar ist.

Für die Auswahl der schließlich zu beurteilenden Krankheitsbilder war wichtig, daß es eine objektive Skala für den Schweregrad von Krankheiten nicht gibt. Als Grundlage der Auswahl wurde eine Statistik der Deutschen Gesetzlichen Rentenversicherung benutzt. Und zwar wurden die Krankheiten nach ihrer Häufigkeit als Ursache einer Berufsunfähigkeit ausgewählt. Die dann weiter ergänzte endgültige Liste bestand aus folgenden zehn nach Möglichkeit umgangssprachlich bezeichneten Krankheiten:

Syphilis
Blinddarmentzündung
Querschnittslähmung
Tuberkulose
Zuckerkrankheit
Krebs
Bluthochdruck
Herzinfarkt
Grippe
Akne.

Diese zehn Krankheiten sollten das "Bezugssystem" bilden für fünf weiter "idealtypisch" gemeinte psychische

Erkrankungen:

Neurose
Alkoholismus
Hirnarteriosklerose ("Verkalkung")
endogene Depression
Schizophrenie.

Als Beurteiler wurden sechs in sich möglichst homogene und voneinander unterschiedenen Gruppen ausgewählt. Homogenität wurde dabei vor allem unter dem Gesichtspunkt spezifischer Beziehung zu Krankheiten bzw. zur Kenntnis über Krankheiten gesehen. Praktisch bedeutsam war dabei auch die Erreichbarkeit dieser Gruppen, womit auf der anderen Seite auf eine empirisch streng repräsentative Auswahl verzichtet werden mußte. Deswegen gebot sich Vorsicht bei der Generalisierung von Ergebnissen. Zum einen sollten Betroffene vertreten sein. Erreicht werden konnten Tuberkulose-Patienten und Alkoholkranke. Unterschiedlichen Informationsstand repräsentierten zwei Gruppen von Medizinstudenten, und zwar einmal eine größere Gruppe vor dem Kursus in Psychiatrie, dann eine in den Personen teilweise identische nach diesem Kursus. Als Vertreter einer Bevölkerungsgruppe ohne wesentliche Krankheitsbelastung dienten zwei Kontrollgruppen, welche beide aus Bundeswehrrekruten bestanden.

Während die Medizinstudenten schriftlich zusätzlich zum jeweils paarweisen Vergleich der Krankheitsbilder ihre Beurteilungskriterien formulierten, erhielten die anderen Gruppen einen informellen Wissenstest über die fünfzehn Krankheitsbilder vorgelegt. Damit konnte der Informationsstand geprüft werden. Zwischen den beiden Gruppen der Rekruten wurde dabei die Reihenfolge der Bearbeitung von Paarvergleichsliste und Wissenstest getauscht. Die zweite Gruppe erhielt zuerst den Wissenstest, welcher als Lerninstrument dienen sollte und deswegen einen geringen Schwierigkeitsgrad hatte.

6.4 Ergebnisse

Aus einer Beurteilung im Paarvergleich durch eine Gruppe von Probanden resultiert eine Skala nach Art einer Thermometerskala. Auf einer solchen Skala sind nicht nur die Rangpositionen, sondern wegen der Intervallskalenqualität auch die Abstände zwischen den Krankheiten interpretierbar. Die psychologische Dimension ist entsprechend der Definition des Beurteilungsaspektes "subjektiv empfundener Schweregrad". Zu beachten ist, daß es auf einer solchen Skala keinen echten Nullpunkt gibt. So kann kein Krankheitsbild in irgendeiner Gruppe identifiziert werden, das als gar nicht schlimm angesehen wird.

Es zeigte sich, daß für alle sechs Gruppen eine Skala der subjektiven Schwere von Krankheiten erstellt werden kann. Diese Skalen sind psychologisch in hohem Maße eindimensional. Sowohl innerhalb der Beurteilergruppen wie auch vor allem zwischen diesen Gruppen überwog die Übereinstimmung in der Beurteilung der Krankheiten weitaus. Unterschiede waren demgegenüber insgesamt nur geringfügig. Das gilt, obwohl die Gruppen sich nach Kenntnis der Krankheiten (Medizinstudenten vs. Laien), nach Bildungsniveau und nach persönlicher Erfahrung mit bestimmten Krankheiten (Patienten vs. Nichtpatienten) teilweise deutlich unterschieden. Auch für Gruppen, welche sich nach Geschlecht und nach Alter unterschieden, ist eine ähnlich große Übereinstimmung festzustellen, wie sekundär eruiert werden konnte.

Für alle Skalen der verschiedenen Gruppen bot sich eine generelle inhaltliche Unterteilung der Krankheiten in Abstufungen des Schweregrades an. Grippe, Akne und Blinddarmentzündung, bei den meisten der sechs Ursprungsgruppen-Skalen auch noch Bluthochdruck, werden als wenig schlimm angesehen. Dann gibt es, deutlich abgehoben von den leichten, einen breiten Bereich mit-

telschwerer bis recht schwerer Erkrankungen: Zuckerkrankheit, Neurose, Alkoholismus, endogene Depression, Syphilis, Tuberkulose, Hirnarteriosklerose (für die Nichtmediziner mit dem erklärenden Zusatz "Verkalkung" versehen). Diese Krankheiten lassen sich in ihrer Erstreckung über die Skalen aber dann nicht mehr deutlich von den als schlimmste bzw. schwerste beurteilten Krankheiten abgrenzen. Bei diesen führt eindeutig als schwerste Krebs, dann folgen Querschnittslähmung und Herzinfarkt, in manchen Beurteilergruppen auch noch Schizophrenie.

Es gibt also offensichtlich, wie aus der guten eindimensionalen Skalierbarkeit gefolgert werden konnte, eine recht einheitliche Vorstellung von Krankheiten und ihrer Schwere. Aus dem Wissenstest konnte belegt werden, daß die inhaltlichen Vorstellungen auch hinreichend zutreffend sind, wenn auch Detailliertheit der Kenntnisse hier nicht zu eruieren war. Das trifft auch auf psychische Krankheiten zu, allerdings bei diesen mit Einschränkungen. Bei der Beurteilung psychischer Krankheiten zeigt sich nämlich eine deutlich größere Unsicherheit des Urteils.

Eine weitere Frage war, ob eine persönliche Erfahrung mit einer bestimmten Krankheit eine andere Einstellung zu dieser Krankheit hervorruft, abzulesen aus einer anderen Einstufung der jeweiligen Krankheit (Alkoholismus und Tuberkulose) auf der Skala dieser Patientengruppen im Vergleich zu den Skalen der übrigen Gruppen. Denkbar war, daß eine solche Krankheit entweder als schlimmer oder auch als weniger schlimm betrachtet wird. Die Gruppen der Tuberkulose-Patienten und der Alkoholiker ordnen aber jeweils Tuberkulose bzw. Alkoholismus nicht wesentlich anders ein als die beiden Kontrollgruppen der Bundeswehrrekruten und die Medizinstudenten.

Die generelle Übereinstimmung zwischen allen Beurtei-

lergruppen im Hinblick auf die Einstufung der verschiedenen Erkrankungen, die man über die Korrelation der Skalen erfassen kann, beruht vor allem auf der sehr übereinstimmenden Skalierung der körperlichen Erkrankungen. In der Reihenfolge der Skalenpositionen der somatischen Krankheitsbilder unterscheiden sich die Gruppen fast nicht. Als Maßzahl für die Übereinstimmung zwischen den Gruppen wurde die mittlere Korrelation zwischen den Skalen allein der somatischen Krankheiten bestimmt; sie beträgt .98. Die Übereinstimmung innerhalb der einzelnen Gruppen in Bezug auf ein bestimmtes somatisches Krankheitsbild ist zu ersehen aus einer Analyse der mittleren Quartilabstände der ursprünglichen Rangplatzverteilungen, welche mit wenigen Ausnahmen für die körperlichen Krankheiten gering sind.

Die Übereinstimmung in der Skalierung und auch in der Anordnung bei den psychiatrischen Krankheitsbildern sowohl innerhalb der Gruppen als auch zwischen den Gruppen ist insgesamt immer noch hoch, jedoch deutlich geringer als bei den somatischen Krankheiten. Von allen Beurteilern werden sie insgesamt weder als leicht noch als sehr schlimm beurteilt. Neurose wird als die leichteste dieser Krankheiten von allen sechs Gruppen angesehen, die anderen vier wechseln jedoch untereinander und in Relation zu den mittelschweren somatischen Krankheiten über die verschiedenen Skalen hinweg immer wieder die Positionen. Die mittlere Korrelation beträgt hier "nur" .85. Für den mittleren Quartilabstand der individuellen Rangwerte pro Beurteilergruppe finden sich höhere Werte als bei den somatischen Krankheiten. Das wurde als Anzeichen dafür gesehen, daß auch innerhalb der Gruppen die Übereinstimmung geringer ist.

Bei einer Unterteilung nach Geschlecht bei den Gruppen, bei denen das möglich war, ergab sich: Weibliche Beurteiler sehen psychiatrische Erkrankungen im allgemeinen

als schlimmer an als die männlichen Mitglieder der gleichen Beurteilergruppe.

Die Kriterien der Beurteilung nach "subjektiv empfundener Schwere" einer Krankheit, so konnte belegt werden, sind im wesentlichen konventionell geprägt: Lebensgefahr, Unheilbarkeit, Beeinträchtigungen in der Lebensführung sind die entscheidenden Gesichtspunkte in allen Gruppen. Hingegen hatten Erwägungen sozialer Folgen einer Erkrankung oder ihr soziales Image keine feststellbare Wirkung auf die Beurteilung. Das gilt auch für die Beurteilung psychischer Erkrankungen.

Es ließ sich weiterhin wahrscheinlich machen, daß eine Verbesserung der Information die Bewertung ändert, weniger bei den somatischen, deutlicher bei den psychiatrischen Krankheiten. Allgemein scheint sich ein höherer Grad von Informiertheit vor allem im Hinblick auf die Differenzierung zwischen den Krankheitsbildern auszuwirken. Es konnte gezeigt werden, daß eine unterschiedlich starke Spreizung der jeweiligen Skalen bei den verschiedenen Gruppen ein Indikator dafür ist. Ohne daß dies eindeutig belegt werden konnte, bot sich als erklärende Hypothese an: Diese Beziehung zwischen "Differenzierung zwischen den Krankheiten" einerseits und "Grad der Kenntnis" andererseits dürfte nicht linear sein. Geringere Kenntnis läßt weniger gut unterscheiden, größere Kenntnis führt zu besserer Differenzierung, noch komplexere Information aber erschwert die Unterscheidung wieder.

Eine Betrachtung definitiver interindividueller Unterschiede in der Beurteilung der psychischen Krankheiten wurde durch den Vergleich von kontrastierenden Gruppen ermöglicht. Diese Kontrastgruppen wurden aus dem Gesamt der vier Gruppen der Nichtmediziner gebildet. Kriterium für die Zuordnung zu einer Kontrastgruppe (z. B. Gruppe mit niedriger Einstufung von Schizophrenie vs. Gruppe mit hoher Einstufung) war

die individuelle extrem niedrige oder extrem hohe rangmäßige Einordnung jeweils eines der fünf psychiatrischen Krankheitsbilder. Ergänzt wurde dieser fünffache Skalenvergleich von je zwei einander entsprechenden Kontrastgruppen durch die Darstellung typischer einzelner Beurteiler. So konnten vor allem weitere Aufschlüsse über die Beweggründe bei der Beurteilung spezifischer Krankheitsbilder gewonnen werden.

Danach besteht eine Art extremer Einschätzung der psychischen Erkrankungen darin, daß sie alle gemeinsam in der jeweiligen Skala nahe bei den leichtesten Erkrankungen eingeordnet werden. Eine solche extreme Einstufung ist aber im wesentlichen durch weniger gute Kenntnis begründet, wie anhand der zugehörigen Kennwerte aus dem Wissenstest belegt werden konnte. Eine Einschätzung psychischer Krankheiten als extrem schlimm aufgrund von relativer Unkenntnis ließ sich hingegen nicht feststellen.

Der Faktor "Kenntnis der Krankheiten" ließ sich durch Herausnahme weniger gut informierter Beurteiler vor einer erneuten Skalierung der Kontrastgruppen-Skalen ausschalten. Es zeigte sich deutlich, daß die insgesamt zehn Kontrastgruppen nicht zehn prinzipiell unterschiedliche Gruppenskalen erstellten. Aufgrund der Übereinstimmung von jeweils fünf der resultierenden Skalen konnte sinnvollerweise nur eine einzige und einheitlich kontrastierende Gruppe charakterisiert werden. Diese Beurteiler ordneten konsistent alle psychischen Krankheiten relativ nahe beieinander an der Spitze der Skalen den schweren Erkrankungen zu. Die Skalen der Gruppen mit jeweils niedriger Einstufung einer der fünf psychiatrischen Erkrankungen gingen hingegen sozusagen in den Skalen der Ursprungsgruppen auf. Sie stimmen mit diesen im wesentlichen überein. Numerisch gezeigt werden konnte das durch die Bestimmung der Korrelationen der Kontrastgruppen-Skalen zu den ursprünglichen sechs Skalen.

Es mußte geschlossen werden, daß bei den sich tatsächlich inhaltlich abhebenden Gruppen eine dezidierte Einstellung zu psychischen Krankheiten anzunehmen ist. Entscheidend ist danach, daß gerade das Moment des Psychischen bei einer Erkrankung ausschlaggebend dafür ist, daß sie als sehr schwer angesehen wird. Einen Gegentyp zu dieser Einstellung, der dadurch charakterisiert ist, daß gerade das psychische Moment als Grund für "nicht schwer" angesehen wird, gab es nur als Einzelfall unter denjenigen, welche über psychische Krankheiten hinreichend gut informiert waren.

Die Bedeutung der Ergebnisse wurde abschließend diskutiert. Es wurde festgestellt, daß sie darin zu sehen ist, daß vor allem die Kenntnisse über eine Krankheit sich als entscheidend für eine Bildung einer Einstellung zu ihr erweisen. Das gilt in besonderem Maße für psychische Erkrankungen. Zugleich aber war darauf zu verweisen, daß generell in einem gewissen Umfang zutreffende Kenntnisse über Krankheiten, auch über psychische Krankheiten, angenommen werden können und daß man diesen objektiven Einschätzungen des Schweregrades - und damit wohl auch der Folgen - als realistisch ansehen kann. Es wurde schließlich die Frage gestellt, ob nicht auf diesem Hintergrund Stil und Inhalte von auf die Öffentlichkeit gezielten Darstellungen über psychische Probleme und Erkrankungen wie auch Selbstdarstellungen aus dem Bereich der Psychologie und Psychiatrie neu überdacht werden müßten. Dabei steht vor allem zur Frage die Neigung von Psychologen und Psychiatern, ihre Probleme und ihre Arbeit für besonders wichtig zu halten, die von ihnen behandelten Erkrankungen als besonders schlimm darzustellen und damit eine eigene Sonderstellung und besondere Ansprüche zu begründen.

Anhang 1:

Literaturverzeichnis

LITERATURVERZEICHNIS

BINDER, J. & ANGST, J.: Soziale Konsequenzen psychischer Störungen in der Bevölkerung, Arch. Psychiatr. Nervenkr. 229 (1981) 355 - 370

BEUTEL, P., KÜFFNER, H. & SCHUBÖ, W.: SPSS 8 - Statistik-Programm-System für die Sozialwissenschaften, 3. Aufl. Stuttgart 1980

BLEULER, E.: Lehrbuch der Psychiatrie, 11. umgearb. Aufl. von M. BLEULER, Berlin 1969

BOCK, R.D.: Remarks on the test of significance for the method of paired comparisons, Psychometrika 23 (1958) 323 - 334

BRAATZ, G.A.: Preference intransitivity as an indicator of cognitive slippage in schizophrenia, J. Abn. Psychol. 75 (1970) 1 - 6

CLAUSS, G. & EBNER, H.: Grundlagen der Statistik, Frankfurt 1971

COOMBS, C.H.: Thurstone's measurement of social values revisited forty years later, J. Person. Soc. Psychol. 6 (1967) 85 - 90

DAVID, H.A.: The method of paired comparisons, London 1963

DAWES, R.M.: Grundlagen der Einstellungsmessung, Weinheim 1967

DIEHL, J.M. & KOHR, H.-U.: Durchführungsanleitungen für statistische Tests, Weinheim 1977

EDWARDS, A.L.: Techniques of attitude scale construction, New York 1957

FEGER, H.: Skalierte Informationsmenge und Eindrucksurteil, Bern 1972

FRÖHLICH, W.D. & BECKER, J.: Forschungsstatistik, 5. Aufl. Bonn 1971

GIGERENZER, G.: Messung und Modellbildung in der Psychologie, München 1981

GUILFORD, J.P.: Psychometric methods, New York 1954

GULLIKSEN, H.A.: A least squares solution for paired comparisons with incomplete data, Psychometrika 21 (1956) 125 - 134

GULLIKSEN, H. & TUKEY, J.W.: Reliability for the law of comparative judgement, Psychometrka 23 (1958) 95 - 110

HERMANN, U.: Medizin-soziologische Analyse der Einflußfaktoren auf "Compliance", Nervenarzt 50 (1979) 102 - 108

HILGARD, E.R., ATKINSON, R.C. & ATKINSON, R.L.: Introduction to psychology, 5. Aufl. New York 1971

HILL, R.J.: A note on inconsistency in paired comparison judgements, Am. Sociol. Revue 18 (1952) 564 - 566

HUBER, G.: Psychiatrie, Stuttgart 1974

JASPERS, K.: Allgemeine Psychopathologie, 7. Aufl. Berlin 1969

JAECKEL, M. & WIESER, S.: Das Bild des Geisteskranken in der Öffentlichkeit, Stuttgart 1970

JENKINS, C.D. & ZYZANSKI, S.J.: Dimensions of belief and feeling concerning three diseases, poliomyelitis, cancer and mental illnes: A factor analytic study Behavioral Science 13 (1968) 372 - 381

KAISER, H.F. & SERLIN, R.C.: Contributions to the method of paired comparisons, Appl. Psychol. Measurement 2 (1978) 421 - 430

KENDALL, M.G.: Rank correlation methods, London 1948

KEUPP, H. (Hrsg.): Der Krankheitsmythos in der Psychopathologie, München 1972

KEUPP, H. (Hrsg.): Normalität und Abweichung, München 1979

KNOPF, A.: Changes in women's opinions about cancer, Soc. Sci. & Medicine 10 (1976) 191 - 195

LEVINE, G.N.: Anxiety about illness: Psychological and social bases J. Health & Human Behavior 3 (1962) 30 - 34

LIENERT, G.A.: Testaufbau und Testanalyse, 3. Aufl. Weinheim 1969

LIENERT, G.A.: Verteilungsfreie Methoden in der Biostatistik,
Band I, Meisenheim 1973
Band II, Meisenheim 1978

LINDEN, M.: Therapeutische Ansätze zur Verbesserung von "Compliance",
Nervenarzt 50 (1979) 109 - 114

MITTENECKER, E.: Planung und statistische Auswertung von Experimenten,
2. Aufl. Wien 1970

MOSTELLER, F.: Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed,
Psychometrika 16 (1951) 207 - 218

MÜLLER- SUUR, H.: Das Psychisch Abnorme - Untersuchungen zur Allgemeinen Psychiatrie,
Berlin 1950

NIE, N.H., HULL, C.H., JENKINS, J.G., STEINBRENNER, K. & BENT, D.H.: SPSS - Statistical package for the social sciences
2. Aufl. New York 1975

ORTH, B.: Einführung in die Theorie des Messens,
Stuttgart 1974

PAULEIKOFF, B. & MESTER, H.: "Einsicht",
in: MÜLLER, C. (Hrsg.), Lexikon der Psychiatrie,
Berlin 1973

QUEKELBERGHE, VAN, R.: Untersuchung zum Selbstkonzept von chronisch Schizophrenen und zum Bild des Geisteskranken,
Psychol. Rundschau 27 (1976) 28 - 39

RABKIN, J.G.: Opinions about mental illness: A review of the literature
Psychol. Bull. 77 (1972) 153 - 171

ROBBINS, P.R.: Some explorations into the nature of anxieties relating to illness,
Gen. Psychol. Mon. 66 (1962) 91 - 141

SIEGEL, S.: Nonparametric statistics for the behavioral sciences
New York 1956

SCHNEIDER, K.: Psychischer Befund und psychiatrische Diagnose,
Leipzig 1939

SCHNEIDER, K.: Die Beurteilung der Zurechnungsfähigkeit, Stuttgart 1948; 4. Aufl. 1961

SCHNEIDER, K.: Psychiatrie heute, Stuttgart 1952; 3. Aufl. 1960

SCHNEIDER, K.: Klinische Psychopathologie, Stuttgart 1950; 8. erg. Aufl. 1967

SIXTL, F.: Meßmethoden der Psychologie, Weinheim 1967; 2. Aufl. 1982

SIXTL, F.: Gedanken über die Verzahnung von Allgemeiner und Differentieller Psychologie, Arch. f. Psychol. 124 (1972) 145 - 157

Statistik der deutschen gesetzlichen Rentenversicherung, Bd. 50: Rentenzugang und Rentenwegfall im Jahre 1977, hrsg. Verband deutscher Rentenversicherer Frankfurt 1978

STATISTISCHES JAHRBUCH 1979, hrsg. Statistisches Bundesamt, Stuttgart/Mainz 1979

SÜLLWOLD, F.: Probleme der Interpretation von Einstellungsskalen, in: HORN, H. (Hrsg.), Psychologie und Pädagogik, Weinheim 1967

SÜLLWOLD, F.: Theorie und Methodik der Einstellungsmessung, in: GRAUMANN, C.F. (Hrsg.), Handbuch der Psychologie, Bd. 7,1 Göttingen 1969 (a)

SÜLLWOLD, F.: Zur Polarität des Ethnozentrismus, in: RUPPERT, J.P.(Hrsg.), Freiheit und Verantwortung in Gesellschaft und Erziehung, Bad Homburg/Berlin/Zürich 1969 (b)

SCHÖNROK, G.: Ein medizinsoziologisches Verfahren zum Erfassen von Meinungen über Krankheiten und deren Images Zs. Ges. Hygiene 24 (1978) 51 - 54

STUMME, W.: Psychische Erkrankungen im Urteil der Bevölkerung München 1975

SZASZ, T.S.: Der Mythos von der seelischen Krankheit (1960); dt. in: KEUPP, H. (Hrsg.), Der Krankheitsmythos in der Psychopathologie, München 1972

THURSTONE, L.L.: A law of comparative judgement, Psychol. Revue 34 (1927) 273 - 286

TORGERSON, W.S.: Theory and methods of scaling,
5. Aufl. New York 1965

WEITBRECHT, H.J.: Psychiatrie im Grundriß,
2. Aufl. Berlin 1968

WIECK, H.H.: Lehrbuch der Psychiatrie,
2. Aufl. Stuttgart 1977

"Wie funktioniert das?" - Der Mensch und seine Krankheiten, hrsg. Fachredaktion für Naturwissenschaften und Medizin des Bibliographischen Instituts,
Mannheim 1973

WILSON, M.E., SAKATA, R. & FRUMKIN, R.M.: Attitudes of some gifted adults, future rehabilitation counselors, and rehabilitation professors toward disabilities,
Psychol. Rep. 22 (1968) 1303 - 1304

WILSON, M.E., BEATTY, J.B. & FRUMKIN, R.M.: Attitudes of future rehabilitation counselors toward eight major disabilities,
Psychol. Rep. 21 (1967) 928

Anhang 2:

Die Paarvergleichslisten

a) Fassung für die Medizinstudenten

b) Fassung für die Patienten- und Kontrollgruppen

Auf den folgenden Seiten finden Sie eine Anzahl von paarweise gegenübergestellten Krankheitsbildern (jedes mit jedem kombiniert), etwa wie folgendes Beispiel:

Bronchitis - Lungenentzündung

Ihre Aufgabe in dieser Untersuchung ist es, bei jedem der vorgestellten Paare zu entscheiden, welches von beiden Krankheitsbildern S i e als das s c h l i m m e r e ansehen.

Hielten Sie also bei dem obigen Beispiel "Lungenentzündung" für s c h l i m m e r als "Bronchitis", dann kreuzen Sie dieses Krankheitsbild des Paares an, und zwar einfach so:

Bronchitis - Lungenentzündung

Hielten Sie dagegen "Bronchitis" für s c h l i m m e r , müßten Sie entsprechend anders ankreuzen:

Bronchitis - Lungenentzündung

Entscheiden Sie Sich aber bitte immer für eine der beiden Alternativen! Die Entscheidung "gleich schlimm" ist Ihnen nicht gestattet!

Manche Bezeichnungen der Krankheitsbilder könnten Ihnen etwas weit gefaßt vorkommen. Es ist aber für diese Untersuchung wichtig, sie so weit gefaßt zu lassen und nicht näher zu differenzieren. Sie sollten dabei jeweils an das übliche Bild denken, so wie Sie es kennen.

Bevor Sie jetzt zügig beginnen, notieren Sie bitte einige persönliche Angaben (die nur Sie als Ihre eigenen identifizieren können!) auf diesem Deckblatt:

Anfangsbuchstabe Ihres Vornamens:	Geburtsdatum: Tag Mon. Jahr	Geschlecht:	Alter:	Wievielte Semester:

Vielen Dank für Ihre Mitarbeit! Beginnen Sie jetzt bitte:

- 1 -

Kreuzen Sie bitte von jedem Paar die Krankheit an,
die Ihrer Meinung nach s c h l i m m e r ist:

Krebs	-	Akne
Grippe	-	Schizophrenie
Neurose	-	Querschnittslähmung
Endogene Depression	-	Blinddarmentzündung
Bluthochdruck	-	Herzinfarkt
Tuberkulose	-	Alkoholismus
Zuckerkrankheit	-	Hirnarteriosklerose
Tuberkulose	-	Syphilis
Zuckerkrankheit	-	Herzinfarkt
Endogene Depression	-	Alkoholismus
Bluthochdruck	-	Querschnittslähmung
Krebs	-	Blinddarmentzündung
Neurose	-	Schizophrenie
Akne	-	Grippe
Endogene Depression	-	Schizophrenie

- 2 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Hirnarteriosklerose	-	Syphilis
Zuckerkrankheit	-	Tuberkulose
Herzinfarkt	-	Alkoholismus
Bluthochdruck	-	Endogene Depression
Querschnittslähmung	-	Blinddarmentzündung
Neurose	-	Krebs
Schizophrenie	-	Akne
Grippe	-	Querschnittslähmung
Endogene Depression	-	Akne
Neurose	-	Herzinfarkt
Tuberkulose	-	Blinddarmentzündung
Bluthochdruck	-	Hirnarteriosklerose
Syphilis	-	Alkoholismus
Krebs	-	Schizophrenie
Bluthochdruck	-	Neurose

- 3 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Grippe	-	Krebs
Querschnittslähmung	-	Akne
Neurose	-	Endogene Depression
Herzinfarkt	-	Blinddarmentzündung
Bluthochdruck	-	Tuberkulose
Hirnarteriosklerose	-	Alkoholismus
Zuckerkrankheit	-	Syphilis
Tuberkulose	-	Hirnarteriosklerose
Syphilis	-	Herzinfarkt
Endogene Depression	-	Zuckerkrankheit
Alkoholismus	-	Querschnittslähmung
Krebs	-	Bluthochdruck
Blinddarmentzündung	-	Schizophrenie
Neurose	-	Grippe
Akne	-	Alkoholismus

- 4 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Zuckerkrankheit	-	Alkoholismus
Bluthochdruck	-	Syphilis
Hirnarteriosklerose	-	Blinddarmentzündung
Neurose	-	Tuberkulose
Herzinfarkt	-	Akne
Grippe	-	Endogene Depression
Schizophrenie	-	Querschnittslähmung
Hirnarteriosklerose	-	Herzinfarkt
Endogene Depression	-	Syphilis
Zuckerkrankheit	-	Querschnittslähmung
Krebs	-	Alkoholismus
Bluthochdruck	-	Schizophrenie
Neurose	-	Akne
Grippe	-	Blinddarmentzündung
Krebs	-	Querschnittslähmung

- 5 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Grippe	-	Herzinfarkt
Tuberkulose	-	Akne
Neurose	-	Hirnarteriosklerose
Syphilis	-	Blinddarmentzündung
Bluthochdruck	-	Zuckerkrankheit
Schizophrenie	-	Herzinfarkt
Endogene Depression	-	Krebs
Schizophrenie	-	Alkoholismus
Zuckerkrankheit	-	Krebs
Querschnittslähmung	-	Syphilis
Hirnarteriosklerose	-	Endogene Depression
Herzinfarkt	-	Tuberkulose
Grippe	-	Bluthochdruck
Blinddarmentzündung	-	Akne
Grippe	-	Zuckerkrankheit

- 6 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Alkoholismus	-	Bluthochdruck
Blinddarmentzündung	-	Zuckerkrankheit
Syphilis	-	Neurose
Akne	-	Hirnarteriosklerose
Tuberkulose	-	Grippe
Schizophrenie	-	Zuckerkrankheit
Syphilis	-	Krebs
Querschnittslähmung	-	Hirnarteriosklerose
Tuberkulose	-	Endogene Depression
Neurose	-	Blinddarmentzündung
Bluthochdruck	-	Akne
Grippe	-	Alkoholismus
Querschnittslähmung	-	Endogene Depression
Herzinfarkt	-	Krebs
Schizophrenie	-	Hirnarteriosklerose

- 7 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist:

Tuberkulose	-	Schizophrenie
Hirnarteriosklerose	-	Grippe
Akne	-	Syphilis
Zuckerkrankheit	-	Neurose
Blinddarmentzündung	-	Alkoholismus
Endogene Depression	-	Herzinfarkt
Tuberkulose	-	Querschnittslähmung
Krebs	-	Hirnarteriosklerose
Syphilis	-	Schizophrenie
Bluthochdruck	-	Blinddarmentzündung
Neurose	-	Alkoholismus
Zuckerkrankheit	-	Akne
Grippe	-	Syphilis
Querschnittslähmung	-	Herzinfarkt
Tuberkulose	-	Krebs

Nachdem Sie jetzt bei den Paaren angegeben haben, welches der jeweiligen Krankheitsbilder Sie als schlimmer ansehen, notieren Sie bitte noch kurz n u r als S t i c h w o r t e die K r i t e r i e n , nach denen Sie entschieden haben, möglichst in der Rangfolge ihrer Wichtigkeit:

1.----------

2.----------

3.----------

4.----------

5.----------

Nochmals Vielen Dank für Ihre Mitarbeit!

Auf den folgenden Seiten finden Sie eine Anzahl von paarweise gegenübergestellten Krankheiten (jede mit jeder kombiniert), etwa wie in folgendem Beispiel:

Bronchitis - Lungenentzündung

Ihre Aufgabe in dieser Untersuchung ist es, bei jedem der Paare zu entscheiden, welche von beiden Krankheiten eines Paares für **S i e p e r s ö n l i c h** die **s c h l i m m e r e** wäre, wenn **S i e** davon **b e t r o f f e n** wären.

Hielten Sie es also etwa bei dem obigen Beispiel für schlimmer, an "Lungenentzündung" zu erkranken als an "Bronchitis", dann müßten Sie diese **s c h l i m m e r e** Krankheit ankreuzen, und zwar einfach so:

Bronchitis - Lungenentzündung X

Hielten Sie dagegen "Bronchitis" für schlimmer, müßten Sie entsprechend anders ankreuzen:

Bronchitis X - Lungenentzündung

Entscheiden Sie sich aber bitte immer für eine der beiden Krankheiten! Die Entscheidung "gleich schlimm" ist Ihnen nicht gestattet! Jedes Blatt enthält eine Reihe von Krankheitspaaren. Achten Sie bitte darauf, daß Sie kein Paar auslassen!

Gehen Sie immer von der Vorstellung aus, Sie selbst wären von der Krankheit betroffen!

Manche Bezeichnungen der Krankheiten könnten Ihnen etwas weit gefaßt vorkommen. Dann sollten Sie dabei jeweils an die übliche Erscheinung der Krankheit denken, so wie Sie sie kennen.

Bevor Sie jetzt zügig beginnen, beantworten Sie bitte noch einige Fragen zu Ihrer Person, indem sie einfach ankreuzen, was zutrifft, bzw. Ihr Lebensalter eintragen:

Wie alt sind Sie? _____ Jahre

Welche Schulbildung haben Sie?

__Hauptschulabschluß
__Abschluß Real-/höhere Schule
__Fachschulabschluß
__Abitur
__ohne Schulabschluß
__Sonstiges

Welche Berufsausbildung haben Sie?

__angelernt
__Lehrberuf
__Studium
__keine Berufsausbildung
__noch in Ausbildung

- 1 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Krebs	-	Akne
Grippe	-	Schizophrenie
Neurose	-	Querschnittslähmung
Endogene Depression	-	Blinddarmentzündung
Bluthochdruck	-	Herzinfarkt
Tuberkulose	-	Alkoholismus
Zuckerkrankheit	-	Hirnarteriosklerose ("Verkalkung")
Tuberkulose	-	Syphilis
Zuckerkrankheit	-	Herzinfarkt
Endogene Depression	-	Alkoholismus
Bluthochdruck	-	Querschnittslähmung
Krebs	-	Blinddarmentzündung
Neurose	-	Schizophrenie
Akne	-	Grippe
Endogene Depression	-	Schizophrenie

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist!, Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Hirnarteriosklerose ("Verkalkung")	-	Syphilis
Zuckerkrankheit	-	Tuberkulose
Herzinfarkt	-	Alkoholismus
Bluthochdruck	-	Endogene Depression
Querschnittslähmung	-	Blinddarmentzündung
Neurose	-	Krebs
Schizophrenie	-	Akne
Grippe	-	Querschnittslähmung
Endogene Depression	-	Akne
Neurose	-	Herzinfarkt
Tuberkulose	-	Blinddarmentzündung
Bluthochdruck	-	Hirnarteriosklerose ("Verkalkung")
Syphilis	-	Alkoholismus
Krebs	-	Schizophrenie
Bluthochdruck	-	Neurose

- 3 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Grippe	-	Krebs
Querschnittslähmung	-	Akne
Neurose	-	Endogene Depression
Herzinfarkt	-	Blinddarmentzündung
Bluthochdruck	-	Tuberkulose
Hirnarteriosklerose ("Verkalkung")	-	Alkoholismus
Zuckerkrankheit	-	Syphilis
Tuberkulose	-	Hirnarteriosklerose ("Verkalkung")
Syphilis	-	Herzinfarkt
Endogene Depression	-	Zuckerkrankheit
Alkoholismus	-	Querschnittslähmung
Krebs	-	Bluthochdruck
Blinddarmentzündung	-	Schizophrenie
Neurose	-	Grippe
Akne	-	Alkoholismus

- 4 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Zuckerkrankheit	-	Alkoholismus
Bluthochdruck	-	Syphilis
Hirnarteriosklerose ("Verkalkung")	-	Blinddarmentzündung
Neurose	-	Tuberkulose
Herzinfarkt	-	Akne
Grippe	-	Endogene Depression
Schizophrenie	-	Querschnittslähmung
Hirnarteriosklerose ("Verkalkung")	-	Herzinfarkt
Endogene Depression	-	Syphilis
Zuckerkrankheit	-	Querschnittslähmung
Krebs	-	Alkoholismus
Bluthochdruck	-	Schizophrenie
Neurose	-	Akne
Grippe	-	Blinddarmentzündung
Krebs	-	Querschnittslähmung

- 5 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Grippe	-	Herzinfarkt
Tuberkulose	-	Akne
Neurose	-	Hirnarteriosklerose ("Verkalkung")
Syphilis	-	Blinddarmentzündung
Bluthochdruck	-	Zuckerkrankheit
Schizophrenie	-	Herzinfarkt
Endogene Depression	-	Krebs
Schizophrenie	-	Alkoholismus
Zuckerkrankheit	-	Krebs
Querschnittslähmung	-	Syphilis
Hirnarteriosklerose ("Verkalkung")	-	Endogene Depression
Herzinfarkt	-	Tuberkulose
Grippe	-	Bluthochdruck
Blinddarmentzündung	-	Akne
Grippe	-	Zuckerkrankheit

- 6 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Alkoholismus	-	Bluthochdruck
Blinddarmentzündung	-	Zuckerkrankheit
Syphilis	-	Neurose
Akne	-	Hirnarteriosklerose ("Verkalkung")
Tuberkulose	-	Grippe
Schizophrenie	-	Zuckerkrankheit
Syphilis	-	Krebs
Querschnittslähmung	-	Hirnarteriosklerose ("Verkalkung")
Tuberkulose	-	Endogene Depression
Neurose	-	Blinddarmentzündung
Bluthochdruck	-	Akne
Grippe	-	Alkoholismus
Querschnittslähmung	-	Endogene Depression
Herzinfarkt	-	Krebs
Schizophrenie	-	Hirnarteriosklerose ("Verkalkung")

- 7 -

Kreuzen Sie bitte jeweils von jedem Paar die Krankheit an, die Ihrer Meinung nach s c h l i m m e r ist! Stellen Sie sich dabei immer vor, Sie seien s e l b s t b e t r o f f e n !

Tuberkulose	-	Schizophrenie
Hirnarteriosklerose ("Verkalkung")	-	Grippe
Akne	-	Syphilis
Zuckerkrankheit	-	Neurose
Blinddarmentzündung	-	Alkoholismus
Endogene Depression	-	Herzinfarkt
Tuberkulose	-	Querschnittslähmung
Krebs	-	Hirnarteriosklerose ("Verkalkung")
Syphilis	-	Schizophrenie
Bluthochdruck	-	Blinddarmentzündung
Neurose	-	Alkoholismus
Zuckerkrankheit	-	Akne
Grippe	-	Syphilis
Querschnittslähmung	-	Herzinfarkt
Tuberkulose	-	Krebs

Anhang 3:

"Wissenstest" mit Schwierigkeitsindices der Items

dazu

a) Instruktion für die Patientengruppen und die Kontrollgruppe 1

b) Instruktion für die Kontrollgruppe 2 ("Wissenstest" vorangestellt)

Bisher haben Sie eine Reihe von Krankheiten daraufhin miteinander verglichen, für wie schlimm Sie diese halten.
Herzlichen Dank für Ihre Mitarbeit bisher!

Nun folgen noch einige Fragen zu den einzelnen Krankheiten, die Sie genauso zügig beantworten können wie die Vergleiche bisher.

Es handelt sich um einige Wissenfragen über die genannten Krankheiten. Wir möchten nur ganz grob feststellen, wie genau Sie über diese Krankheiten informiert sind.

Damit die Beantwortung für Sie recht einfach ist, sind bei jeder Krankheit fünf Feststellungen angeführt. Drei von diesen fünf Feststellungen sind jeweils richtig, die zwei anderen sind falsch für die genannte Krankheit.
Ihre Aufgabe ist es einfach, die richtigen Feststellungen herauszufinden und anzukreuzen.

Ein Beispiel, wie Sie es machen sollen:

Wenn die Krankheit, nach der gefragt wird, "Bronchitis" wäre, könnte das so aussehen:

B r o n c h i t i s

... ist eine Krankheit,die erblich ist .....O
... ist eine chronische Entzündung der oberen Luftwege ..X
... kann oft Folge gewohnheitsmäßigen Rauchens sein X
... und jetzt käme die 4. Feststellung .....O
... und jetzt noch eine 5. Feststellung X

Sie sehen, im Beispiel sinddie 2. und die 3. Feststellung angekreuzt, denn sie sind richtig, sie treffen für Bronchitis zu. Die 1. Feststellung ist für Bronchitis falsch, deswegen nicht angekreuzt.
Die 4. Feststellung wäre falsch, deswegen nicht angekreuzt, die 5. wiederum richtig (insgesamt drei sind immer richtig!), deswegen ist sie angekreuzt.

Nun können Sie umblättern und anfangen!

Auf den folgenden Seiten finden Sie einige Wissenfragen über einzelne Krankheiten. Wir möchten nur ganz grob feststellen, wie genau Sie über diese Krankheiten informiert sind.

Damit die Beantwortung für Sie recht einfach ist, sind bei jeder Krankheit fünf Feststellungen angeführt. Drei von diesen fünf Feststellungen sind jeweils richtig, die zwei anderen sind falsch für die genannte Krankheit.
Ihre Aufgabe ist es einfach, die richtigen Feststellungen herauszufinden und anzukreuzen.
Ein Beispiel, wie Sie es machen sollen:

Wenn die Krankheit, nach der gefragt wird, "Bronchitis" wäre, könnte das so aussehen:

B r o n c h i t i s

... ist eine Krankheit, die erblich ist 0

... ist eine chronische Entzündung der oberen Luftwege.....X
... kann oft Folge gewohnheitsmäßigen Rauchens sein.....X
... und jetzt käme die 4.Feststellung 0
... und jetzt noch eine 5. Feststellung.....X

Sie sehen, im Beispiel sind die 2. und die 3. Feststellung angekreuzt, denn sie sind richtig, sie treffen für Bronchitis zu. Die 1. Feststellung ist für Bronchitis falsch, deswegen nicht angekreuzt.
Die 4. Feststellung wäre falsch, deswegen nicht angekreuzt, die 5. wiederum richtig (insgesamt drei sind immer richtig!), deswegen ist sie angekreuzt.

Nun können Sie umblättern und anfangen!

Die Untersuchung wird mit Genehmigung des
Bundesministers der Verteidigung durchgeführt.
Sie ist unter der Nr. 2/1a1/80 registriert.

K1 K2 TB Alk

A k n e

9o 97 92 93
... tritt als Hauterkrankung mit Pickeln vor allem in der Jugendzeit auf, heilt aber meist ohne Folgen aus . . . . . ⊗

95 93 83 1oo
... ist belastend, weil sie das Aussehen des Betroffenen unvorteilhaft verändert, ist aber nicht gefährlich . . . . . ⊗

93 97 92 93
... führt bei falscher Behandlung zur Vernarbung der Haut, dann vor allem im Gesicht . . . . . ⊗

93 94 9o 98
... kann jedes Organ befallen und geht mit körperlichem Verfall einher . . . ○

95 94 77 88
... ist eine seelische Erkrankung, die sich in einer Unfähigkeit zu sachentsprechendem Reagieren in oft alltäglichen Lebenssituationen äußert. . ○

B l u t h o c h d r u c k

1oo 99 98 1oo
... ist eine Erkältungskrankheit, die durch Infektion (Ansteckung) entsteht ○

82 75 79 78
... ist eine Krankheit, bei der der Kranke bemerkt, daß er sein Verhalten oder sein Empfinden nicht vernünftig steuern kann, und darunter leidet. ○

98 96 92 93
... ist eine Erkrankung, bei der der Blutdruck dauerhaft stark erhöht ist . ⊗

82 93 92 93
... wenn er längere Zeit besteht, kann er gefährliche körperliche Folgen haben . . . . . ⊗

89 82 92 93
... ist eine Erkrankung, die häufig langdauernde Behandlung erfordert . . . ⊗

A l k o h o l i s m u s

92 94 9o 98
... verändert auf die Dauer den Kranken in seiner Persönlichkeit (seinem Wesen) und verursacht oft schwere körperliche und geistige Beeinträchtigungen (z.B. Wahnvorstellungen) . . . . . ⊗

97 96 9o 1oo
... ist eine Erkrankung des Stoffwechsels im Zuckerhaushalt und kann Bewußtlosigkeit verursachen ("Überzucker"). . . . . ○

97 94 85 1oo
... entsteht oft durch Gewöhnung an übermäßigen Alkoholkonsum, der dann in seinem Ausmaß vom Erkrankten nicht mehr gesteuert werden kann. . . . ⊗

89 87 69 9o
... kann heute mit Medikamenten behandelt werden, die Behandlung ist aber langwierig . . . . . ○

93 92 88 95
... ist nur durch eine zumeist langwierige Behandlung mit psychologischen Mitteln (Psychotherapie) und absolutes Vermeiden von Alkoholgenuß zu heilen . . . . . ⊗

K1 K2 TB Alk

Herzinfarkt

87 92 88 85
... ist eine akut auftretende Erkrankung des Herzmuskels, die Folge einer Verstopfung der Herzkranzgefäße (= Blutgefäße am Herzen) und ist im akuten Stadium zumeist sehr schmerzhaft . . . . . ⊗

75 82 63 71
... nur, wenn er nicht erkannt oder nicht behandelt wird, ist er gefährlich . ○

98 94 94 98
... ist eine körperliche Erkrankung, deren Ursachen noch nicht völlig geklärt sind, und bei der bösartige Geschwulste entstehen, die Tochtergeschwulste in anderen Körperteilen bilden können . . . . . ○

87 9o 81 88
... ist im akuten Stadium lebensgefährlich, aber bei richtiger und ausreichender Behandlung bestehen gute Heilungsaussichten . . . . . ⊗

98 97 9o 98
... verlangt immer eine Umstellung der Lebensgewohnheiten, z.B. Verzicht auf Rauchen, Kontrolle von Übergewicht, Vermeidung von zu großer körperlicher und seelischer Belastung . . . . . ⊗

Schizophrenie

9o 72 71 8o
... ist vor allem eine Alterskrankheit, deren Folge eine geringere Durchblutung des Gehirns ist. . . . . ○

95 92 85 9o
... ist eine seelische Erkrankung (oft "Geisteskrankheit" genannt), die den Erkrankten in seiner Persönlichkeit (seinem Wesen) sehr verändern kann . . . . . ⊗

93 85 85 93
... geht oft mit großen Angstgefühlen oder auch mit Wahnvorstellungen einher und beeinträchtigt die geistige Leistungsfähigkeit und zwischenmenschliche Beziehungen . . . . . ⊗

98 99 9o 95
... ist eine Krankheit, die durch Infektion entsteht und vor allem die Lunge befällt und langdauernd verläuft . . . . . ○

95 92 85 93
... wird heute vor allem mit Medikamenten, aber gleichzeitig auch mit psychologischen Mitteln (Psychotherapie) behandelt, die Behandlung ist aber zumeist langwierig und manchmal erfolglos . . . . . ⊗

Blinddarmentzündung

95 82 98 9o
... ist nur gefährlich, wenn sie nicht erkannt oder nicht behandelt wird. . ⊗

98 96 1oo 98
... ist eine Erkältungskrankheit, die durch Infektion (Ansteckung) entsteht ○

89 79 88 88
... ist mit bestimmten Medikamenten (z.B. Penicillin) gut zu behandeln, wenn sie zeitig erkannt wird. . . . . ○

95 96 96 98
... geht mit heftigen körperlichen Beschwerden, mit Schmerzen im Bauch und Fieber einher . . . . . ⊗

98 97 98 98
... ist durch eine einfache Operation gewöhnlich ziemlich sicher zu behandeln und zu heilen . . . . . ⊗

K1	K2	TB	Alk	
				Krebs
1oo	99	96	95	... ist eine körperliche Erkrankung, deren Ursachen noch nicht völlig geklärt sind, und bei der bösartige Geschwulste entstehen, die Tochtergeschwulste in anderen Körperteilen bilden können. . . . . ⊗
1oo	97	1oo	1oo	... ist eine Geschlechtskrankheit, die durch Infektion (Ansteckung) übertragen wird und sehr oft mit Hautausschlägen einhergeht . . . . . ○
1oo	99	98	98	... ist eine Krankheit, die oft zu spät erkannt wird und dann gewöhnlich lebensgefährlich ist, und jedes Organ befallen kann . . . . . ⊗
98	99	9o	95	... ist nur durch eine zumeist lange Behandlung mit psychologischen Mitteln (Psychotherapie) und absolutes Vermeiden von Alkoholgenuß zu heilen . . ○
98	99	92	93	... eine Behandlung dieser Krankheit durch Medikamente, Operation, Bestrahlung bringt häufig nur eine Verlängerung des Lebens, keine Heilung . . . . . ⊗
				Neurose
93	83	75	83	... ist eine seelische Erkrankung, die sich in einer Unfähigkeit zu sachentsprechendem Reagieren in oft alltäglichen Lebenssituationen äußert . . . . . ⊗
93	86	88	85	... ist eine Krankheit, bei der der Kranke bemerkt, daß er sein Verhalten oder sein Empfinden nicht "vernünftig" steuern kann, und darunter leidet . . . . . ⊗
92	82	88	85	... beruht auf einer Schädigung des Rückenmarks, sehr häufig nach Unfallverletzung der Wirbelsäule, die zu einer völligen Lähmung und Empfindungslosigkeit der tiefer liegenden Körperteile führt . . . . . ○
98	85	77	85	... ist mit einem bestimmten Medikament (z.B. Penicillin) gut zu behandeln,wenn sie zeitig erkannt wird . . . . . ○
98	87	85	95	... ist eine Krankheit, die durch Behandlung mit psychologischen Mitteln (Psychotherapie) geheilt oder zumindest gebessert werden kann, wenn auch die Behandlung manchmal schwierig ist . . . . . ⊗
				Grippe
1oo	99	96	95	... ist eine Erkältungskrankheit, die durch Infektion (Ansteckung) entsteht . ⊗
98	97	96	1oo	... geht oft mit großen Angstgefühlen oder auch mit Wahnvorstellungen einher und beeinträchtigt die geistige Leistungsfähigkeit und zwischenmenschliche Beziehung . . . . . ○
87	87	9o	93	... ist eine Erkrankung, bei der sich der Betroffene bei normalem Verlauf mit Fieber, Gliederschmerzen und Halsweh krank fühlt, die aber unbehandelt nach wenigen Tagen folgenlos wieder ausheilt . . . . . ⊗
7o	52	71	71	... gefährdet bei einer Ansteckung eigentlich nur kleine Kinder und alte Leute ⊗
56	59	73	76	...unbehandelt oder nicht richtig behandelt sind schwere Folgeerkrankungen möglich . . . . . ○

K1 K2 TB Alk

Hirnarteriosklerose ("Verkalkung")

85 85 73 88
... ist im akuten Stadium lebensgefährlich, aber bei richtiger und ausreichender Behandlung bestehen gute Heilungsaussichten . . . . . ○

75 89 79 85
... bewirkt eine Hemmung aller Gefühle und Antriebe und treibt häufig zu Selbstmordgedanken oder auch Selbstmordversuchen . . . . . ○

89 99 85 93
... ist vor allem eine Alterskrankheit, deren Folge eine geringere Durchblutung des Gehirns ist . . . . . ⊗

98 94 88 93
... beeinträchtigt die geistige Leistungsfähigkeit, vor allem das Gedächtnis, und kann das Wesen eines Menschen erheblich verändern . . . . . ⊗

74 75 75 88
... ist durch Behandlung mit Medikamenten zu lindern, aber nicht völlig zu heilen . . . . . ⊗

Zuckerkrankheit

97 89 9o 98
... ist eine Erkrankung des Stoffwechsels im Zuckerhaushalt und kann Bewußtlosigkeit verursachen ("Überzucker") . . . . . ⊗

98 97 98 98
... macht die Einhaltung einer Diät (Bestimmte Nahrungsmittel dürfen garnicht, die meisten nur in ganz bestimmten Mengen genossen werden) notwendig und verlangt deswegen dauernde Selbstkontrolle . . . . . ⊗

1oo 94 9o 98
... ist eine Krankheit, die durch Behandlung mit psychologischen Mitteln (Psychotherapie) geheilt oder zumindest gebessert werden kann, wenn auch die Behandlung manchmal schwierig ist . . . . . ○

98 97 92 1oo
... erfordert zumeist eine Dauerbehandlung mit Medikamenten, unter Umständen tägliche Spritzen; unbehandelt oder nicht richtig behandelt macht sie schwere Folgeerkrankungen möglich . . . . . ⊗

95 89 92 98
... die betroffenen Körperteile, zu denen die Nervenbahnen unterbrochen sind, können ihre normale Tätigkeit nicht mehr ausüben . . . . . ○

Endogene Depression

9o 77 6o 9o
... verlangt immer eine Umstellung der Lebensgewohnheiten, z.B. Verzicht auf Rauchen, Kontrolle von Übergewicht, Vermeidung von zu großer körperlicher und seelischer Belastung . . . . . ○

95 89 88 98
... ist eine seelische Erkrankung (oft "Gemütskrankheit" genannt), welche die Stimmung des Erkrankten zu tiefer Traurigkeit und Verzweiflung verändert, ohne daß ein wirklicher Anlaß vorhanden ist . . . . . ⊗

89 89 83 1oo
... ist sehr oft unheilbar, wenn auch nicht tödlich, und führt häufig zu einem Leben im Rollstuhl . . . . . ○

98 94 83 93
... bewirkt eine Hemmung aller Gefühle und Antriebe und treibt häufig zu Selbstmordgedanken oder auch Selbstmordversuchen . . . . . ⊗

82 76 71 93
... wird vor allem durch Medikamente, begleitend auch mit psychologischen Mitteln (Psychotherapie) behandelt und heilt zumeist nach kürzerer oder längerer Zeit vollkommen aus . . . . . ⊗

K1 K2 TB Alk

S y p h i l i s

98 97 98 100
... ist eine Geschlechtskrankheit, die durch Infektion (Ansteckung) übertragen wird und sehr oft mit Hautausschlägen einhergeht . . . . . ⊗

85 89 92 93
... eine Behandlung durch Medikamente, Operation, Bestrahlung bringt häufig nur eine Verlängerung des Lebens, keine Heilung . . . . . ○

93 94 92 100
... ist mit bestimmten Medikamenten (z.B. Penicillin) gut zu behandeln, wenn sie zeitig erkannt wird . . . . . ⊗

97 99 94 98
... ist eine seelische Erkrankung, die sich in einer Unfähigkeit zu sachentsprechendem Reagieren in oft alltäglichen Lebenssituationen äußert . . ○

90 94 90 90
... ist eine Krankheit, die nur gefährlich wird, wenn sie bis ins letzte Stadium kommt, weil sie dann Gehirn und Nervensystem angreift und zur "Verblödung" führen kann . . . . . ⊗

Q u e r s c h n i t t s l ä h m u n g

100 97 100 98
... beruht auf einer Schädigung des Rückenmarks, sehr häufig nach Unfallverletzungen der Wirbelsäule, die zu einer völligen Lähmung und Empfindungslosigkeit der tiefer liegenden Körperteile führt . . . . . ⊗

98 97 94 100
... ist sehr oft unheilbar, wenn auch nicht tödlich, und führt häufig zu einem Leben im Rollstuhl . . . . . ⊗

100 99 96 98
... ist durch eine einfache Operation ziemlich sicher zu behandeln und zu heilen . . . . . ○

98 96 96 98
... bei der Krankheit können die betroffenen Körperteile, zu denen die Nervenbahnen unterbrochen sind, ihre normalen Tätigkeiten nicht mehr ausüben . . . . . ⊗

97 92 96 98
... beeinträchtigt die geistige Leistungsfähigkeit, vor allem das Gedächtnis, und kann das Wesen eines Menschen erheblich verändern . . . . . ○

T u b e r k u l o s e

97 87 94 100
... ist eine seelische Erkrankung, die sich in einer Unfähigkeit zu sachentsprechendem Reagieren in oft alltäglichen Lebenssituationen äußert . . . ○

100 92 96 100
... ist eine Krankheit, die durch Infektion (Ansteckung) entsteht und vor allem die Lunge befällt und zumeist langdauernd verläuft . . . . . ⊗

93 93 100 98
... kann heute mit Medikamenten behandelt werden, die Behandlung ist aber langwierig . . . . . ⊗

90 73 85 90
... heilt in den allermeisten Fällen aus, kann aber Dauerschäden zurücklassen ⊗

92 75 90 90
... ist belastend, weil sie das Aussehen des Betroffenen unvorteilhaft verändert, aber nicht gefährlich . . . . . ○

V i e l e n D a n k f ü r I h r e M i t a r b e i t !

Anhang 4:

Verzeichnis der beteiligten Gruppen und Kliniken bzw. Institutionen

Gruppen der Medizinstudenten (Med 1 und Med 2):

Studentinnen und Studenten im zumeist 4. klin. Semester zu Beginn und Ende der Vorlesung von Prof. Dr. med. K. Wanke, Medizinische Fakultät der Universität des Saarlandes, 6650 Homburg/Saar

Tuberkulose- Patienten (Tb):

1. Caritasklinik - Abt. für Pneumonologie
   Chefarzt Dr. med. Strunz
   6990 Bad Mergentheim
2. Sonnenhof-Sanatorium
   Chefarzt Dr. med. Kläser
   7988 Wangen/Allgäu
3. Fachklinik Senne I
   Chefarzt Dr. med. Lutz
   4800 Bielefeld
4. Klinik Heidehaus
   Chefarzt Dr. med. Mall
   3000 Hannover 21

Alkohol-Patienten (Alk):

1. Fachklinik Mahlertshof
   Ltg. Dipl.-Soz.arb. Rudeck u. Dr. med. Klippel
   6419 Burghaun/Krs. Hünfeld
2. Fachklinik Schloß Mackenzell
   Ltg. Frau Dipl.-Soz.arb. Flügel u. Dr. med. Klippel
   6418 Hünfeld

Kontrollgruppen (Kontr 1 und Kontr 2)

Soldaten der Bundeswehr aus zwei Ausbildungkompanien des HSK 16, Nieder-Auerbach-Kaserne,
6630 Zweibrücken/Pfalz
mit freundlicher Genehmigung des Bundesministers der Verteidigung (Nr. 2/101/80)

Lebenslauf

Am 16. September 1943 wurde ich als Sohn des Kaufm. Angestellten Hans Wollny und seiner Frau Karola geb. Kimmel in Wasserburg/Inn als deutscher Staatsangehöriger geboren.

Ab April 1950 besuchte ich die Volksschule, ab April 1954 das Mathematisch-Naturwissenschaftliche Gymnasium in Herten (Westfalen). Von April 1958 an besuchte ich das Humanistische Gymnasium Collegium Augustinianum in Gaesdonck bei Goch, das ich nach bestandener Reifeprüfung im Februar 1963 verließ.

Von 1963 bis 1971 studierte ich Philosophie und Theologie, zuerst für ein Semester an der Universität Freiburg, dann an der Philosophisch-Theologischen Hochschule Albertus Magnus-Akademie in Walberberg bei Bonn. Dies Studium schloß ich im Juni 1971 mit dem Grad eines Lektors der Theologie ab. Nach einem Semester Soziologie studierte ich von 1972 bis 1976 Psychologie an der Johann Wolfgang Goethe-Universität zu Frankfurt am Main. Im Juli 1976 erwarb ich dort das Diplom in Psychologie.

Am 1. September 1976 erhielt ich eine Anstellung als Diplom-Psychologe beim Frankfurter Verein für Soziale Heimstätten e.V., Frankfurt. Vom 1. November 1976 bis zum 30. April 1979 arbeitete ich als Klinischer Psychologe am Psychiatrischen Landeskrankenhaus Philippshospital in Riedstadt-Goddelau (Hessen). Seit dem 1. Mai 1979 bin ich als Klinischer Psychologe an der Universitäts-Nervenklinik - Psychiatrie - in Homburg/Saar tätig.

Meine akademischen Lehrer in meinem Studium der Psychologie an der Universität Frankfurt waren die Professoren H.J. Bochnik, A. Degenhardt, I. Deusinger, P. Kutter, V. Lange, A. Mitscherlich, H.K. Müller, V. Sarris, F. Süllwold, L. Süllwold, L. Tent, H.M. Trautner, H. Vogel, K. Wanke, F.K. Wender, J. Zoltobrocki.

Herrn Prof. Dr. F. Süllwold, meinem akademischen Lehrer in Frankfurt, fühle ich mich zu besonderem Dank verpflichtet. Beim Entstehen dieser Arbeit hat er mich an wichtigen Punkten kritisch und wegweisend beraten und mich über lange Zeit hin wohlwollend begleitet.

Mein Dank gilt ebenso Herrn Prof. Dr. K. Wanke, Direktor der Universitäts - Nervenklinik, - Psychiatrie -, in Homburg/Saar. Er hat mir die zur Erstellung der Arbeit erforderlichen wissenschaftlichen Arbeitsmittel großzügig zur Verfügung gestellt und mich jederzeit unterstützt.

Nicht vergessen werden sollen schließlich alle im Anhang genannten beteiligten Institutionen und ihre verantwortlichen Leiter, vor allem aber auch die beteiligten Probanden, welche durch ihre freiwillige Mithilfe die Grundlage für diese Untersuchung gaben.